穆维新 主编

张慎武 袁 浩 副主编

数据路由与交换技术

Data Routing and Switching Technology

清华大学出版社

VISIT…

LANZAROTE
Caliente.COM

数据路由与交换技术

穆维新　主编
张慎武　袁　浩　副主编

清华大学出版社
北　京

内 容 简 介

本书围绕数据通信，对路由和交换技术进行了全面的阐述。全书共分为12章，包含OSI七层参考模型和IP/TCP协议栈、数据网通信和互联网技术，MAC帧交换和IP分组路由、局域网设计和VLAN配置、各种数据链路层协议原理及配置、各种路由协议的原理及配置、子网划分和网络地址转换、网络操作系统及服务器设置等。本书理论联系实际，可以使读者学以致用。

本书内容丰富，技术新颖，论述清晰，配置实例众多，主要面向网络相关专业技术人员，可作为高等院校电子信息、计算机网络、电子商务等专业的高年级教材或学习参考用书。

本书封面贴有清华大学出版社防伪标签，无标签者不得销售。
版权所有，侵权必究。侵权举报电话：010-62782989　13701121933

图书在版编目(CIP)数据

数据路由与交换技术/穆维新主编；张慎武，袁浩副主编．—北京：清华大学出版社，2018
ISBN 978-7-302-50558-7

Ⅰ．①数…　Ⅱ．①穆…　②张…　③袁…　Ⅲ．①数据通信—通信技术　Ⅳ．①TN919

中国版本图书馆CIP数据核字(2018)第142118号

责任编辑：王　芳
封面设计：常雪影
责任校对：李建庄
责任印制：董　瑾

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦A座　　**邮　　编**：100084
社 总 机：010-62770175　　**邮　　购**：010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
课件下载：http://www.tup.com.cn，010-62795954
印 装 者：三河市铭诚印务有限公司
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张**：19.25　　**字　　数**：470千字
版　　次：2018年12月第1版　　**印　　次**：2018年12月第1次印刷
定　　价：59.00元

产品编号：076737-01

前　言

本书围绕数据通信，阐述了路由协议、交换原理等基础理论，并结合互联网结构对路由交换设备的常用命令做了具体介绍，给出了各种配置实例。本书跟踪网络新技术，理论联系实际，使读者能融会贯通，提高学习的实用性。

在学习本书的过程中，建议读者尽量多去思考一些问题。比如，一个终端设备产生的数据信息需要经过什么样的承载网来传送？数据信息在网络中如何完成交换？网络如何确定路由并将数据信息送到目的地？组成数据网的有关设备与OSI七层参考模型是什么样的对应关系？局域网如何规划？网络运行的有关协议以及它的作用是什么？协议地址和物理地址在网络中起的作用是什么？如何区分子网和VLAN？如何配置交换路由设备？等等，既是读者需要了解的，也是本书要详尽介绍的。

全书共分12章。第1章至3章，介绍OSI参考模型、TCP/IP协议栈和互联网技术，包含系统指标、传输网、传输接口和介质，以及从对物理层到应用层的深入探讨，是对以前所学网络知识的巩固和总结；第4章介绍数据通信网，包含数字数据网(DDN)、分组交换网(X.25)、帧中继(FR)和异步传输模式(ATM)；第5章、第6章主要介绍路由器、交换机的工作原理及其配置；第7章介绍链路层技术及配置，包含高级数据链路控制规程(HDLC)、点到点协议(PPP)和以太网(Ethernet)；第8章是数据网络配置，主要介绍访问控制列表(ACL)、网络地址转换(NAT)，以及X.25和FR的配置；第9章、第10章介绍动态路由协议及配置，包含自治区域内部协议：路由信息协议(RIP)、开放最短路由优先协议(OSPF)和中间系统到中间系统协议(IS-IS)，以及自治区域外部协议(BGP)；第11章介绍网络操作系统(NOS)与服务器的具体设置；第12章介绍局域网，包含虚拟局域网(VLAN)在内的设计及配置。

本书每个章节相对有一定的独立性。作为教材时可结合以前开设课程的相关内容，选择有关章节系统地学习，建议电子信息类、计算机类和电子商务等专业为48学时。参加工作的读者可以根据自己的实际情况和兴趣有重点地学习。

书中有关路由器、交换机等设备的讲述，主要采用Cisco和华为的产品及其模拟器，使读者尽可能多了解一些不同设备的命令及配置。

中原工学院张慎武完成了本书第9～12章编写；郑州大学西亚斯国际学院袁浩完成了本书第6～8章的编写；郑州工业应用技术学院穆维新完成了其余章节的编写并对全书进行统稿。在本书的编写过程中，参考了有关作者的文献和网络咨询信息，采用了有关网络设备制造商的技术资料，郑州大学的师生给予了热情的支持，在此一并表示感谢！

由于时间仓促，书中难免有不足之处，敬请读者批评指正。

编　者

2018年6月

目 录

第 1 章 Internet 基础 …… 1

1.1 Internet 通信技术释义 …… 1
1.2 OSI 参考模型 …… 4
1.2.1 OSI-RM 功能 …… 4
1.2.2 OSI-RM 通信 …… 6
1.3 数据通信系统及指标 …… 9
1.3.1 数据通信系统 …… 9
1.3.2 系统性能指标 …… 12
1.4 数据交换技术 …… 17
1.5 数据传输网技术 …… 21
1.5.1 SDH …… 21
1.5.2 WDM …… 24
1.5.3 PTN …… 25
1.5.4 承载 IP 传输网演进技术 …… 27
1.6 网络接口与介质 …… 28
1.6.1 广域网接口与介质 …… 28
1.6.2 以太网接口与介质 …… 30
习题 …… 33

第 2 章 TCP/IP 协议栈 …… 35

2.1 TCP/IP 协议栈概述 …… 35
2.2 网络接口层 …… 36
2.2.1 物理层 …… 36
2.2.2 数据链路层 …… 38
2.3 网络层 …… 41
2.3.1 网络层路由及其功能 …… 41
2.3.2 IP …… 42
2.3.3 ICMP 和 IGMP …… 46
2.3.4 ARP 和 RARP …… 48
2.4 传输层 …… 49
2.4.1 传输层协议功能与端口 …… 49
2.4.2 TCP …… 50
2.4.3 UDP 和 SCTP …… 54
2.5 应用层 …… 56

习题 …… 58

第 3 章 Internet …… 60

3.1 Internet 设备 …… 60
3.1.1 网络设备概念模型 …… 60
3.1.2 网络设备 …… 61
3.2 IP 地址 …… 64
3.2.1 子网划分 …… 64
3.2.2 变长子网掩码 …… 69
3.2.3 无类域间路由 …… 70
3.3 Internet 应用协议 …… 71
3.3.1 DNS(域名服务器) …… 71
3.3.2 FTP(文件传输协议) …… 74
3.3.3 Telnet(远程登录协议) …… 75
3.3.4 SMTP(简单邮件传送协议) …… 76
3.3.5 WWW(万维网) …… 76
3.3.6 DHCP(动态主机配置协议) …… 78
3.3.7 SNMP(简单网络管理协议) …… 79
3.4 Internet 接入技术 …… 80
3.4.1 CHINANET 与接入技术概述 …… 80
3.4.2 基于协议的接入 …… 81
3.4.3 以专线方式和电话拨号接入 …… 83
3.4.4 ADSL 宽带接入 …… 84
3.4.5 混合网络和无源光网络 …… 85
3.4.6 综合业务接入网 …… 86
习题 …… 87

第 4 章 数据通信网 …… 89

4.1 数字数据网 …… 89
4.1.1 DDN 组成 …… 89
4.1.2 DDN 用户接入方式及其应用 …… 92
4.2 公共交换分组数据网 …… 93
4.2.1 X.25 交换 …… 94
4.2.2 用户接入分组网 …… 97
4.3 帧中继 …… 98
4.3.1 FR 交换 …… 98
4.3.2 FR 网络 …… 102
4.4 异步传输模式(ATM) …… 104
4.4.1 ATM 交换 …… 104

4.4.2 ATM 网络 …… 108
习题 …… 110

第 5 章 路由器 …… 112

5.1 路由器技术 …… 112
5.1.1 路由器基本概念 …… 112
5.1.2 路由器构成 …… 114
5.1.3 路由器工作原理 …… 117
5.2 路由器配置 …… 120
5.2.1 路由器启用 …… 120
5.2.2 路由器配置模式 …… 123
5.2.3 路由器基本配置 …… 125
5.2.4 静态路由的配置 …… 131
5.3 路由器配置实例 …… 134
5.3.1 多路由器组网及配置 …… 134
5.3.2 测试路由器配置 …… 136
习题 …… 137

第 6 章 交换机 …… 138

6.1 以太网交换机 …… 138
6.1.1 交换技术 …… 138
6.1.2 数据帧转发与网段划分 …… 140
6.1.3 交换机互连 …… 143
6.2 多层交换机 …… 144
6.2.1 交换技术 …… 144
6.2.2 交换原理 …… 145
6.3 交换机配置 …… 149
6.3.1 常用配置命令 …… 149
6.3.2 配置实例 …… 152
习题 …… 155

第 7 章 数据链路层协议 …… 156

7.1 高级数据链路控制规程 …… 156
7.1.1 HDLC 技术 …… 156
7.1.2 HDLC 配置 …… 158
7.2 点到点协议 …… 159
7.2.1 PPP 技术 …… 159
7.2.2 PPP 配置 …… 164
7.3 以太网 …… 168

7.3.1 以太网技术 …… 168
7.3.2 以太网接口 …… 171
7.3.3 以太网接口配置 …… 176
习题 …… 178

第 8 章 数据网络配置 …… 179

8.1 访问控制列表 …… 179
8.1.1 ACL 概述 …… 179
8.1.2 ACL 配置 …… 180
8.2 网络地址转换 …… 184
8.2.1 NAT 简述 …… 184
8.2.2 NAT 配置命令 …… 184
8.2.3 NAT 配置实例 …… 187
8.3 分组交换(X.25)配置 …… 192
8.3.1 X.25 基本配置 …… 192
8.3.2 X.25 典型配置举例 …… 195
8.4 帧中继配置 …… 196
8.4.1 FR 配置命令 …… 196
8.4.2 FR 典型配置实例 …… 197
习题 …… 200

第 9 章 自治区域内部协议 …… 201

9.1 路由信息协议 …… 201
9.1.1 RIP 报文结构 …… 201
9.1.2 RIP 工作原理 …… 202
9.1.3 RIP 工作流程 …… 204
9.1.4 RIP 基本配置 …… 205
9.1.5 RIP 配置实例 …… 207
9.2 开放最短路由优先协议 …… 211
9.2.1 OSPF 报文交换 …… 211
9.2.2 OSPF 路由 …… 215
9.2.3 单区域 OSPF 配置实例 …… 218
9.2.4 多区域 OSPF 配置实例 …… 222
9.2.5 OSPF 其他配置 …… 224
9.3 中间系统到中间系统协议 …… 228
9.3.1 IS-IS 工作原理 …… 228
9.3.2 IS-IS 配置 …… 229
习题 …… 231

第 10 章 自治区域外部网关协议 …… 232
10.1 BGP 工作原理 …… 232
10.1.1 BGP 路由 …… 232
10.1.2 BGP 报文和状态机 …… 233
10.2 BGP 配置 …… 236
10.2.1 BGP 基本配置 …… 236
10.2.2 BGP 配置实例 …… 238
习题 …… 244

第 11 章 NOS 与服务器设置 …… 245
11.1 网络操作系统 …… 245
11.1.1 操作系统简介 …… 245
11.1.2 Windows 操作系统 …… 247
11.1.3 其他操作系统 …… 248
11.2 网络操作系统与 TCP/IP …… 249
11.2.1 应用程序接口软件 API …… 249
11.2.2 Socket 基本函数 …… 250
11.2.3 C/S 构架下的 Socket 通信 …… 253
11.3 Windows Server 2016 安装 …… 256
11.3.1 在 VirtualBox 中配置虚拟机 …… 256
11.3.2 Windows Server 2016 的安装 …… 257
11.4 Windows 服务器配置 …… 262
11.4.1 服务器角色选择 …… 263
11.4.2 Web 服务器配置 …… 265
11.4.3 DNS 服务器配置 …… 266
11.4.4 DHCP 服务器配置 …… 268
11.4.5 FTP 服务器配置 …… 269
习题 …… 270

第 12 章 VLAN 与局域网 …… 272
12.1 VLAN 协议及其技术 …… 272
12.1.1 VLAN 协议 …… 272
12.1.2 VLAN 链路 …… 273
12.1.3 VLAN 划分 …… 275
12.1.4 VLAN 路由 …… 277
12.2 VLAN 配置 …… 279
12.2.1 VLAN 基本配置 …… 279
12.2.2 VLAN 路由配置 …… 280

12.3 VLAN 互连 …… 281
12.3.1 RIP 实现 VLAN 互连 …… 281
12.3.2 OSPF 实现 VLAN 互连 …… 283
12.4 局域网设计与配置 …… 287
12.4.1 局域网模型 …… 287
12.4.2 局域网设计 …… 289
12.4.3 局域网配置 …… 292
习题 …… 298

第 1 章 Internet 基础

国际标准化组织(International Organization for Standardization,ISO)推出了描述网络工作机理的开放系统互连参考模型(Open System Interconnection Reference Model,OSI-RM),使 Internet 的相关标准受益匪浅。本章首先介绍数据通信及计算机网络常用术语,然后对 Internet 基础内容进行叙述:OSI-RM 各层功能、协议及通信过程;数据通信系统模型、构成及其各种性能指标;报文交换、分组交换、软交换等各种数据交换技术;SDH、WDM 和 PTN 以及 IP 传输网演进技术;广域网、以太网的传输接口和介质等。

1.1 Internet 通信技术释义

数据通信及计算机网络的一些常用术语介绍如下。

1. 数据(data)

数据是对客观事实进行描述与记载的物理符号。它是信息的载体,可以是数字、文字、语言、图形和图像等,可以分为模拟数据和数字数据两种。现代通信网络中所谓的数据,是相对于传统的基于电路交换网的语音,是侧重于基于分组网中的信息,如数据单元。

2. 信息(information)

信息是数据的集合、含义与解释。数据是相对具体的概念,信息是相对抽象的概念,两者相对存在,有时可以将二者等同起来,一般都体现在开放系统互连参考模型高层,如第七层收发的信息。

3. 信号(signal)

信号是数据或信息传输过程中的表现形式,也可以理解成信号是数据的电编码,如电信号等。信号可以分为模拟信号与数字信号两种,一般都体现在 OSI-RM 低层,如第一层收发的信号。

4. 模拟信号(analogous signal)

代表信号的取值是连续的,是随时间连续变化的信号,相位和幅值都是连续的。

5. 数字信号(digital signal)

代表信号的取值是离散的,在时间上是离散的信号,在幅度上是经过量化后,仅包含有限数目的信号值,最常见的是二值信号。

6. 消息(message)

消息通常指要传递的一个完整的内容,如在通信中发送一条信令就称为一个消息,在软件运行中的一条程序就是一个消息。消息在不同的环境中可以理解为启示、通知、报文等,也可以认为消息是信息的表现形式,信息是消息的具体内容。

7. 信道(channel)

信道一般是用来表示传送信息的通路。一般从数据链路层两端看进去的通路为逻辑信道,而从物理层两端看进去的通路为物理信道,物理信道又分为模拟信道和数字信道。

8. **报文**(message)

报文是网络中交换与传输的数据单元,也是网络传输的单元。报文包含了将要发送的完整数据信息以及报文头,其长度各不一致。报文在传输过程中会被系统不断地封装成分组或数据包以及帧来传输。

9. **分组**(packet)

分组是在网络中传输的二进制格式的单元,为了提高通信的性能和可靠性,每个用户发送的数据会被系统分成多个更小的部分,并在每个部分的前面加上一些必要的控制信息组成的首部,构成了一个分组。如处于 OSI-RM 第三层的 X.25 协议就是分组格式。

10. **数据包**(data packet)

数据包通常指的是 TCP/IP 协议通信传输中的数据单元,简称为"包"。数据包是指其本身包含了足够的寻址信息,可独立地从源主机传输到目的主机。如处于 OSI-RM 第三层的 IP 协议就是数据包格式,也称 IP 包。

11. **数据报**(data gram)

数据报是面向无连接的数据传输,其工作过程类似于报文交换。采用数据报方式传输时,被传输的分组称为数据报。如处于 OSI-RM 第四层的 UDP 协议就属于数据报文格式。

12. **数据段**(data segment)

数据段是面向连接的数据传输,其工作过程是将上层进入该层的报文可以分成更小的数据段,再加入包含段序号在内的段头,然后进行逐层下交传送。如处于 OSI-RM 第四层的 TCP 协议就是数据段格式。

13. **帧**(frame)

帧是数据链路层的传输单元。它将上层传入的数据包添加一个头部和尾部,组成了帧。如以太网的 MAC 协议子层,包含数据包、MAC 地址等,形成了 MAC 帧格式。

14. **客户**(client)**和服务器**(server)

其是指通信中所涉及的两个应用进程。客户是服务的请求方,服务器是服务的提供方。

15. **基带传输**(baseband transmission)

基带传输又称数字传输,是指把要传输的数据转换为数字信号,使用固定的频率在信道上传输。基带传输是一种最基本的传输方式,是典型的矩形电脉冲信号,其频谱包括直流、低频和高频等多种成分。在数据通信中,由计算机发出的二进制数字信号形式称为方波,人们把这种方波固有的频带称为基带。在基带传输中,整个信道只传输一种信号,通信信道利用率低。一般来说,发送端通过编码器将信源的数据变换为直接传输的数字基带信号,在接收端由译码器进行解码,恢复发端数据。基带传输不需要调制解调器,适合短距离的数据传输,如一个企业的局域网就可以采用这种方式将计算机连接到一起。

数据编码方法有单极性不归零码、双极性不归零码、单极性归零码、曼彻斯特码和差分曼彻斯特码等。最后两种编码不含直流分量,包含时钟脉冲,便于双方自同步,应用较多。

16. **频带传输**(frequency band transmission)

频带传输又称模拟传输,是指信号以正弦波形式传播。频带传输就是先将基带信号变换(调制)成便于在模拟信道中传输的、具有较高频率范围的模拟信号(称为频带信号),再将这种频带信号在模拟信道中传输。计算机网络的远距离通信通常采用的是频带传输。用基带信号对载波波形的某些参量进行控制,使这些参量随基带矩形电脉冲变化,这就是调制。

已调信号通过线路传输到接收端,然后经过解调恢复为原始基带矩形电脉冲。这种频带传输不仅克服了目前许多长途线路不能直接传输基带信号的缺点,而且能实现多路复用的目的,从而提高了通信线路的利用率。当然频带传输在发送端和接收端都要设置调制解调器。

数字信号转换为模拟信号的三种技术是频移键控(Frequency-Shift Keying,FSK)、幅度键控(Amplitude Shift Keying,ASK)和移相键控(Phase Shift Keying,PSK)。

17. 宽带传输(wide band transmission)

在局域网中,存在基带传输和宽带传输两种方式,基带传输数据速率低于宽带传输。一个宽带信道可以被划分为多个逻辑基带信道。宽带传输能把声音、图像、数据等信息综合到一个物理信道上进行传输。宽带传输采用频带传输技术,但频带传输不一定是宽带传输。通过借助频带传输,可以将链路容量分解成两个或更多的信道,每个信道可以携带不同的信号,这就是宽带传输。宽带传输中的所有信道都可以同时发送信号,如 CATV 等。

18. 基本速率接口(BRI)

BRI 是 ISDN(综合数字业务网)采用的 2B+D 接口,其中 B 是速率为 64kbps 的数字信道,D 是速率为 16kbps 的数字信道。

19. 基群速率接口(PRI)

PRI 是 ISDN 采用的 30B+D,也称一次群速率接口,其中 B 和 D 的速率均为 64kbps 的数字信道。B 信道主要用于传送用户信息流;D 信道主要用于传送电路交换的信令信息,也用于传送分组交换的数据信息。

20. 网络协议(network protocol)

网络协议是指网络上所有设备,包含服务器、计算机、交换机、路由器、防火墙等之间通信规则的集合,它规定了通信时信息必须采用的格式和这些格式的意义。大多数网络都采用分层的体系结构,在网络的各层中存在着许多协议。网络协议使网络上各种设备能够相互交换信息,如常见的 TCP/IP 协议栈。网络协议主要由以下三个要素组成:语法,即数据与控制信息的结构或格式;语义,即需要发出何种控制信息,完成何种动作以及做出何种响应;同步,即事件实现顺序的详细说明。

21. 广域网(Wide Area Network,WAN)

它的作用范围通常为几十到几千千米,是 Internet 的核心部分,其任务是通过长距离运送主机所发送的数据。一般连接广域网各交换机的都是高速链路,具有较大的通信容量。

22. 城域网(Metropolitan Area Network,MAN)

其作用范围在广域网和局域网之间,约为 5~100km。其传输速率一般在 100Mbps 以上。城域网可以为一个或几个单位所拥有,但是也可以是一种公用通信网,用来将多个局域网进行互连。目前,很多城域网均采用以太网技术,因此城域网有时也常纳入局域网的范围。

23. 局域网(Local Area Network,LAN)

一般通过专用高速通信线路把许多台计算机连接起来,速率一般在 10Mbps 以上,但在地理上则局限在较小的范围。在局域网的初期,一个单位往往只有一个局域网,但现在局域网已经非常广泛,一个学校都拥有许多个互连局域网,通常称为校园网。

24. 个人区域网(Personal Area Network,PAN)

个人区域网就是在个人工作的地方把属于个人使用的电子设备用无线技术连接起来的

网络，因此也通常称为无线个人区域网(Wireless PAN，WPAN)，其范围大约为10m。

顺便指出，若中央处理机之间的距离非常近(如1m的数量级或甚至更小些)，则一般就称之为多处理机系统而不称它为计算机网络。

25．接口(interface)

这里指计算机、集线器、交换机、路由器等设备连接其他网络设备的硬件接口(hardware interface)，也称物理接口(physical interface)，而非逻辑接口(logical interface)，如RJ45、f1/0、Serial0/1、E1等。在网络设备配置中还会用到子接口(subinterface)，它是一种逻辑接口，它和物理接口在使用上没有大的区别。

物理端口(physical port)(有时简称端口)也称接口，注意实际区别，不要混淆。

26．端口(port)

这里指计算机或交换机、路由器等网络内部连接的逻辑端口(logical port)，也称虚拟端口(virtual port)，或协议端口(protocol port)。例如传输层TCP、UDP包含的以下端口：

熟知端口(familiar port)或周知端口(well-known ports)，端口号为0～1023，其中80端口分配给WWW服务，21端口分配给FTP服务等；

注册端口(registered port)号为1024～49151，分配给用户进程或应用程序；

动态端口(dynamic ports)号为49152～65535，动态分配给某种服务。

1.2 OSI参考模型

在数据通信逐步占据通信领域主导地位的过程中，各厂商纷纷推出自己的协议，由于多种协议的并存，使得网络兼容已成为主要问题。因此，使得ISO早在1984年就提出的OSI-RM成为数据通信的基础模型，Internet的TCP/IP协议栈与OSI参考模型的关系是密不可分的。

1.2.1 OSI-RM功能

国际标准化组织负责制定大型网络的标准，包括与Internet相关的标准。由它提出的OSI-RM描述了网络的工作机理，为数据通信网络构建了一个易于理解的模型。OSI七层参考模型及功能如图1.1所示，它保证了不同网络设备之间的兼容性和互操作性。OSI参考模型的第五层到第七层称为高层(upper layer)，又叫主机层(host layer)，而在计算机网络的TCP/IP协议栈中只体现为一层，统一称为应用层，有时也称为用户层(user layer)或业务层(business layer)。

在实际应用中，有关网络层次划分的原则如下：

(1) 各层功能明确。即每一层的划分都应有自己明确的、与其他层不同的基本功能。当某一层的具体实现方法或功能发生变化时，只要接口不变，就不会对其他相邻层产生影响。

(2) 层间接口清晰，应尽量减少跨越接口的通信量。

(3) 层数适中，层数应足够多，以避免不同的功能混杂在同一层中，但也不能太多，否则体系结构会过于庞大，增加各层服务的开销。

(4) 网络中各节点具有相同或不同层次，但不同节点的对等层应具有相同的功能。

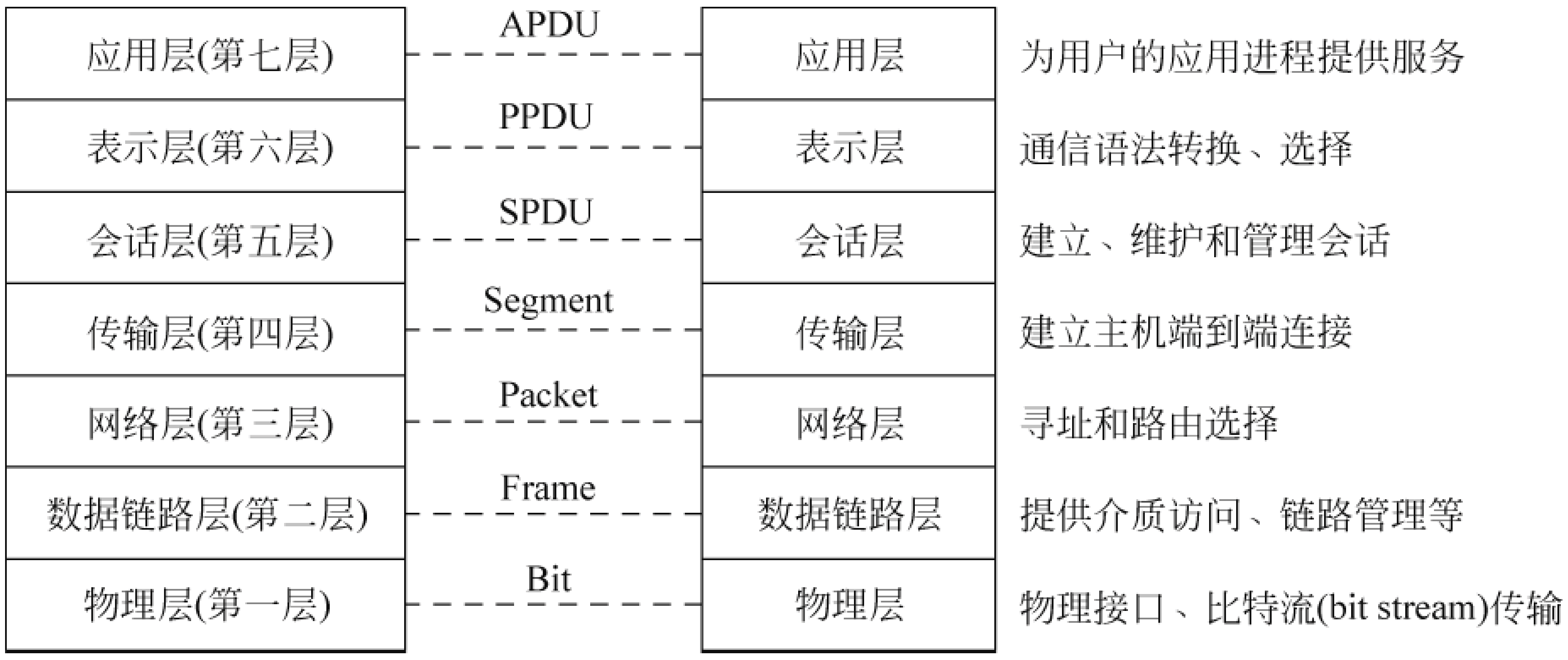


图 1.1 OSI 七层参考模型及功能

OSI 参考模型第一层到第四层称为低层(lower layer),又叫介质层(media layer),主要负责数据在网络中的传送,网络互连设备往往位于最下面的三层,也就是计算机网络的通信子网。低层通常以硬件和软件相结合的方式来实现。高层用于保障数据的正确传输,通常以软件方式来实现。各层的主要功能和协议介绍如下。

(1) 物理层。第一层是物理层(physical layer),它涉及在通信信道(channel)上传输的原始比特流,实现传输数据所需要的机械、电气、功能特性及过程等手段。物理层涉及电压、电缆线、数据传输速率、接口等的定义。

物理层的典型规范有 EIA/TIA RS-232、EIA/TIA RS-499、V.35、RJ-45 等。

(2) 数据链路层。第二层是数据链路层(data link layer),它的主要功能是负责数据链路信息从源点传输到目的点的数据传输与控制,保证将源端主机网络层的数据包准确无误地传送到目的主机的网络层。数据链路层的帧使用物理层提供的比特流传输服务来到达目的主机数据链路层。为了保证数据传输的准确无误,数据链路层还负责网络拓扑、差错校验、流量控制等。流量调控可以在数据链路层实现,也可以由传输层实现。数据链路层传输的基本单位是帧。数据链路层与物理地址、网络拓扑、线缆规划、错误校验和流量控制等有关。

数据链路层的代表协议有 SDLC、HDLC、PPP、STP、帧中继等。

(3) 网络层。第三层是网络层(network layer),它可以通过路由选择协议来确定数据包从源端到目的端路由。网络层的主要功能是建立、维护和拆除网络连接、组包/拆包、路由选择和拥塞控制等。数据链路层协议只是两个直接连接节点间的通信协议,它不能解决数据经过通信子网中多个转接节点的通信问题,而网络层的主要目的就是要为报文分组以最佳路径通过通信子网到达目的主机提供服务,网络层不关心网络的拓扑结构与所使用的通信介质。

网络层有代表性的协议有 IP、IPX、ICMP(Internet Control Message Protocol,因特网控制消息协议)、IGMP(Internet Group Management Protocol,因特网组管理协议)、ARP(Address Resolution Protocol,地址解析协议)和 RARP(Reverse Address Resolution Protocol,反向地址解析协议)等。

(4) 传输层。第四层是传输层(transport layer),它的主要设置目的就是在源主机和目的主机进程之间提供可靠的端到端通信。基本功能是从会话层接收数据,并且在必要的时

候把它分成较小的单元,传递给网络层,并确保到达对方的各段信息正确无误。传输层是用户资源子网与通信子网的桥梁,主要功能为:连接管理;负责传输连接的建立、维护与释放,传输连接的建立过程称为“握手”;流量控制;传输层在发送本层数据分组时,还要确保数据的完整性;差错检测与恢复;提供用户要求的服务质量和端到端的可靠通信。

常见的传输层协议有 TCP、UDP、SCTP、SPX 等。

(5) 会话层。第五层是会话层(session layer),它通过执行多种机制在应用程序间建立、维持和终止会话。会话层机制包括计费、话路控制、会话参数协商等。会话层在协调不同应用程序之间的通信时要涉及会话层,该层使每个应用程序知道其他应用程序的状态。所谓会话,是指在两个用户之间为交换信息而按照某种规则建立的一次暂时联系。同时,会话层也提供双工(duplex)协商、会话同步等;会话层提供远程会话地址、会话建立后的管理,提供把报文分组重新组成新报文的功能。

常见的会话层协议有结构化查询语言(Structured Query Language,SQL)、网络文件系统(Network File System,NFS)、远程过程调用(Remote Procedure Call,RPC)、网络基本输入/输出系统(Network Basic Input/Output System,NetBIOS)、Windows 系统等。

(6) 表示层。第六层是表示层(presentation layer),它提供数据格式转换服务、数据加密、数据表示标准服务等。表示层为应用层提供服务,该服务层处理的是通信双方之间的数据表示问题,如所传输信息的语法和意义,它把来自应用层与计算机有关的数据格式处理成与计算机无关的格式,以保障对端设备能够准确无误地理解发送端数据。表示层的主要功能为语法转换、传送语法的选择、常规功能等。

常见的表示层协议有数据结构标准的 EBCDIC(Extended Binary Coded Decimal Interchange Code)、ASCII(American Standard Code for Information Interchange);图像标准的 JPEG(Joint Photographic Experts Group)、TIFF(Tagged Image File Format)、GIF;视频标准的 MIDI(Musical Instrument Digital Interface)、MPEG(Moving Picture Experts Group)、QuickTime 等。

(7) 应用层。第七层是应用层(application layer),它是 OSI 参考模型最靠近用户的一层,它为用户的应用进程访问 OSI 环境提供服务。应用层识别并验证目的通信方的可用性,使协同工作的应用程序之间同步。

应用层的代表协议有 Telnet、FTP、HTTP、SNMP 等。

1.2.2 OSI-RM 通信

1. OSI 参考模型数据通信过程

可以将终端或网络设备之间平行相对应的层称为对等层或对应层。将终端或网络设备中上、下相接的层称为相邻层。对等层之间通过相同的协议通信,相邻层之间联系通过内部原语。因此,对等层传输的信息就称为协议数据单元(Protocol Data Unit,PDU)。相应地,应用层数据称为应用层协议数据单元(Application Protocol Data Unit,APDU),表示层数据称为表示层协议数据单元(Presentation Protocol Data Unit,PPDU),会话层数据称为会话层协议数据单元(Session Protocol Data Unit,SPDU)。而习惯上又将传输层数据称为段(segment),网络层数据称为数据包(packet),数据链路层数据称为帧(frame),物理层数据称为比特流。

例如，两个终端设备的对等传输层利用数据段进行通信，传输层的段成为网络层数据包的一部分，网络层数据包又成为数据链路层帧的一部分，最后转换成比特流传送到对端物理层，又依次到达对端物理层、数据链路层、网络层、传输层，实现了对等层之间的通信。

如果我们形象地理解 OSI-RM，就是在高层形成一个裸体信息，在发端的每层下行过程，裸体信息是在穿衣服，每层都因各种原因需要加一件衣服，到收端以后就是脱衣服的一个上行过程，发端对等层穿上什么衣服，收端对等层就要脱掉什么衣服，由于对应层执行的是同一种协议，才能保证衣服不会被脱错，通过穿、脱衣服这个过程，就可以明白通信双方的真实意图，以便执行下一步的行动。对于为什么要穿这么多的衣服，就相当于人们要经过千山万水，必须要穿上不同的衣服才能最终到达目的地。数据信息也是一样，要经过错综复杂的网络及设备，就必须要执行不同的协议才能最终到达目的地。我们将穿衣服的过程称为封装，将脱衣服的过程称为解封装，这一点在以后的学习中会得到更加深入的理解，也就是说每一层都有它存在的必要性。数据包经过中间网络设备的过程可以理解为在更换衣服。

关于上、下层的相邻层关系，我们还可以将它理解为用户和服务的关系，即上层是下层的用户，下层是为上层提供服务。每一层利用下一层提供的服务，使用自己的协议与对等层通信。也就是说，终端或网络设备的每一层并不能直接与对端相对应层直接通信，而是通过下一层为其提供的服务来间接与对端对等层相互传递数据，每一层使用各自的协议，但对等层的协议必须相同，以保证对等层之间能够准确无误地传递数据。例如，应用层协议 E-mail 应用程序不会和对端应用层 Telnet 应用程序通信，但可以和对端 E-mail 应用程序通信。下一层通过服务访问点(Service Access Point，SAP)为上一层提供服务。

封装(encapsulation)是指网络节点(node)将要传送的数据用特定的协议头打包来传送数据，有时候，我们也可能在数据尾部加上报文，这些就是封装过程。封装就是以保证数据能够准确无误地到达目的地，被对端设备理解、执行。

【例 1】 用 OSI 七层模型的概念，概述数据从主机 A 到达主机 B 的传送过程。

答：首先，主机的高层将信息转化为能够在网络中传播的应用层协议数据单元(APDU)；如不考虑表示层、会话层，数据就直接下交到传输层，加上传输层报头，形成段；再下交到网络层，加上网络层报头，形成数据包；继续下交到数据链路层加上数据链路层报头形成帧；最后送到主机接口所在的物理层，将数据帧转换为比特流，然后经物理层形成比特流进入通信网络。

在通信网络数据传输的路径中可能要经过多个路由器。那么，每个路由器的物理层在收到比特流后，往上送至数据链路层，由数据链路层从比特流中取出帧，再从帧中提取 IP 数据报上交网络层。路由器的网络层根据 IP 数据报的首部信息，找出转发路由后再将 IP 数据报下送至数据链路层，重新封装成新的帧，然后交给物理层发送给下一个路由器。

依此类推，被传输的数据最终到达主机 B 的物理层，依次解封装。由主机 B 的数据链路层取出 IP 数据报，再经网络层、传输层到达应用层，最后应用层将数据交给主机 B 的应用进程。

2. 面向连接服务与无连接服务

OSI 参考模型以及其他协议栈提供的服务可以分为两种方式，即面向连接的服务和无连接的服务，而这两种服务的具体实现可以是虚电路(或虚链路)服务和数据报服务。

虚电路服务：在虚电路服务方式中，为了进行数据的传输，网络的源主机和目的主机之

间先要建立一条逻辑通道。虚电路服务方式是网络层向传输层提供的一种使所有分组能按顺序到达目的主机的可靠的数据传送方式。也可以理解为在进行数据交换的两端主机之间存在着一条为它们提供服务的虚电路。

数据报服务：数据报服务类似于邮政系统的信件投递。每个分组都携带完整的源、目的节点的地址信息，独立地进行传输，每当经过一个中间节点时，都要根据目标地址和网络当前的状态，按一定的路由选择算法选择一条最佳的输出线，直至传输到目的节点。

（1）面向连接服务（connect-oriented service）。面向连接服务是指在使用该服务之前用户首先要建立连接，而在使用完服务之后，用户应该释放连接，当被叫用户拒绝连接时，连接宣告失败。它适用于延迟敏感性应用。

在建立连接阶段，在有关的服务原语以及协议数据单元中，必须给出源主机和目的主机的地址，建立虚电路连接；在数据传输阶段，可以使用一个连接标识符来表示上述这种连接关系。

通常面向连接服务是可靠的报文序列服务，从来不丢失数据（可靠的服务是由接收方确认收到的每一份报文，使发送方确信它发送的报文已经到达目的地这一方法来实现的，确认过程增加了额外的开销和延迟。通常这也是值得的，但有时也不尽然）。在建立连接之后，每个用户可以发送可变长度（在某一限度之内）的报文，这些报文按顺序发送给远端的用户，用户对这些报文的接收也是有顺序的。面向连接的服务比较适用于在一定期间内向同一个目的地发送很多报文的情况，对于发送很短的零星报文，面向连接的服务显得开销过大。

（2）无连接服务（connectionless service）。无连接服务是指两个实体之间的通信不需要先建立好一个连接，因此其下层的有关资源不需要事先进行预定保留，这些资源是在数据传输时动态地进行分配的，它适用于延迟不敏感的应用。

无连接服务就好比邮政系统，每个报文（信件）带有完整的目的地址，并且每一个报文都独立于其他报文，经由系统选定的路线传递。在正常情况下，当两个报文发往同一目的地时，先发的先收到。但是，也有可能先发的报文在途中延误了，后发的报文反而先收到。无连接服务的特征是它不需要通信的两个实体同时处于激活状态，而只需要正在工作的实体处于激活状态。它的优点是灵活方便和比较迅速，但无连接服务不能防止报文的丢失、重复或失序。因此它比较适合传送少量的零星的报文。

并不是所有的应用程序都需要连接。例如，电子邮件越来越普及，电子邮件发送者可能不希望仅为了发一条消息而去经历建立和拆除连接的麻烦。百分之百的可靠性也没有必要；这里所需要的仅是发送一个报文，只要到达的可能性很大就行了，不需要保证一定收到。对于一些允许延迟的应用程序，如文字处理等，往往也使用无连接的服务。

目前，网络层协议通常只提供无连接的服务，不保证数据包的有序、可靠地传输，而数据可靠传输功能通常放在传输层或数据链路层实现。

3. 端到端连接过程

在 OSI 参考模型中，多个应用程序可以共享同一个传输连接，称为多路复用（multiplex）。多路复用是指多个应用程序共享同一个传输层建立的连接进行数据的传送。传输层把上层发来的不同应用程序数据分成段，按照先到先发（FIFO）的原则（或者其他原则）发送数据段。这些数据段可以去往同一目的地，也可以去往不同目的地。

接着要说明的是终端和服务器在传输层的端到端通信过程，如终端开始调用服务器的

WWW 应用程序时，服务器软件为每一个应用程序设置一个端口号(port number)，此端口号与网络设备物理端口不同，是一个应用程序或协议的虚拟接口，在传输层生成数据段。接下来，服务器的传输层必须和终端主机建立一条端到端的会话虚连接。为了能够开始数据传输，服务器和终端主机的两个应用程序通知各自的操作系统开始初始化连接。两个操作系统上的协议模块在网络上互相发送报文通信，以保证双方虚连接建立。当双方所有同步工作完成之后，端到端虚连接便已建立，数据传输开始。在传输过程中，服务器和终端主机继续以它们的协议软件进行通信，以验证数据是否正确接收。当终端设备收到数据流时，它对这些数据流进行分离和排序，以使传输层能够正确地将数据流送到终端主机。当数据传输结束后，双方协商断开虚连接。

4. 流量控制

在数据传输过程中，由于带宽、各种各样网络设备速率不匹配导致的延迟等，网络也有可能在某一节点产生拥塞导致数据包丢失等，所以要实施流量控制。目前有三种常用的流量控制技术：缓存(buffering)、源抑制报文(source quench messages)和窗口机制(windowing)。

网络设备使用缓存技术把内存中暂时不能处理的突发性数据存放在缓冲区内，待网络设备空闲时再发送。缓存技术可以初步解决数据拥塞问题。然而，当网络中数据流量持续增多时，缓冲区有可能也会过载，从而使得缓冲区数据溢出，导致数据丢失。这时，可以使用源抑制技术来降低网络流量。数据接收端设备通过向源端发送源抑制报文，请求源端降低数据发送速率，防止网络过载。源端设备发送一定数量数据包后，要求目的端返回确认信息，这就是所谓的窗口机制，网络可以通过这种窗口机制方式来实现流量控制。窗口机制因为使用了肯定确认(positive acknowledgement)技术，通常被认为是非常可靠的。

肯定确认技术的工作原理为：当目的设备接收到源设备发送的数据包时，向源端发送确认报文，源设备收到确认报文后，继续发送数据包，如此重复；当源设备发送数据包后没有收到确认报文，在一定时间后(源设备在发送数据包时启动计时器，即计时器计时结束的时间)，源设备降低数据传输速率，重发数据包。

1.3 数据通信系统及指标

数据通信系统是通过数据电路将分布在远地的数据终端设备与计算机系统连接起来，实现数据传输、交换、存储和处理的系统。数据通信系统的基本构成如图 1.2 所示。

1.3.1 数据通信系统

1. 数据通信系统的模型

典型的数据通信系统的模型如图 1.2(a)所示，可以划分为三大部分，分别是源系统、传输系统和目的系统。计算机网络是数据通信的一个应用领域，数据通信系统的模型同样适用于计算机网络。

(1) 源系统一般由信源和发送器系统两部分组成。

信源主要负责把用户需要发送的信息内容进行采集和变换，形成计算机系统能够识别和传输的数据。如语音信号经过处理后存储到计算机中，产生传输中需要的数据比特流。

发送器系统主要功能有两部分：其一是把源系统提供的数据比特流变换成有一定规律

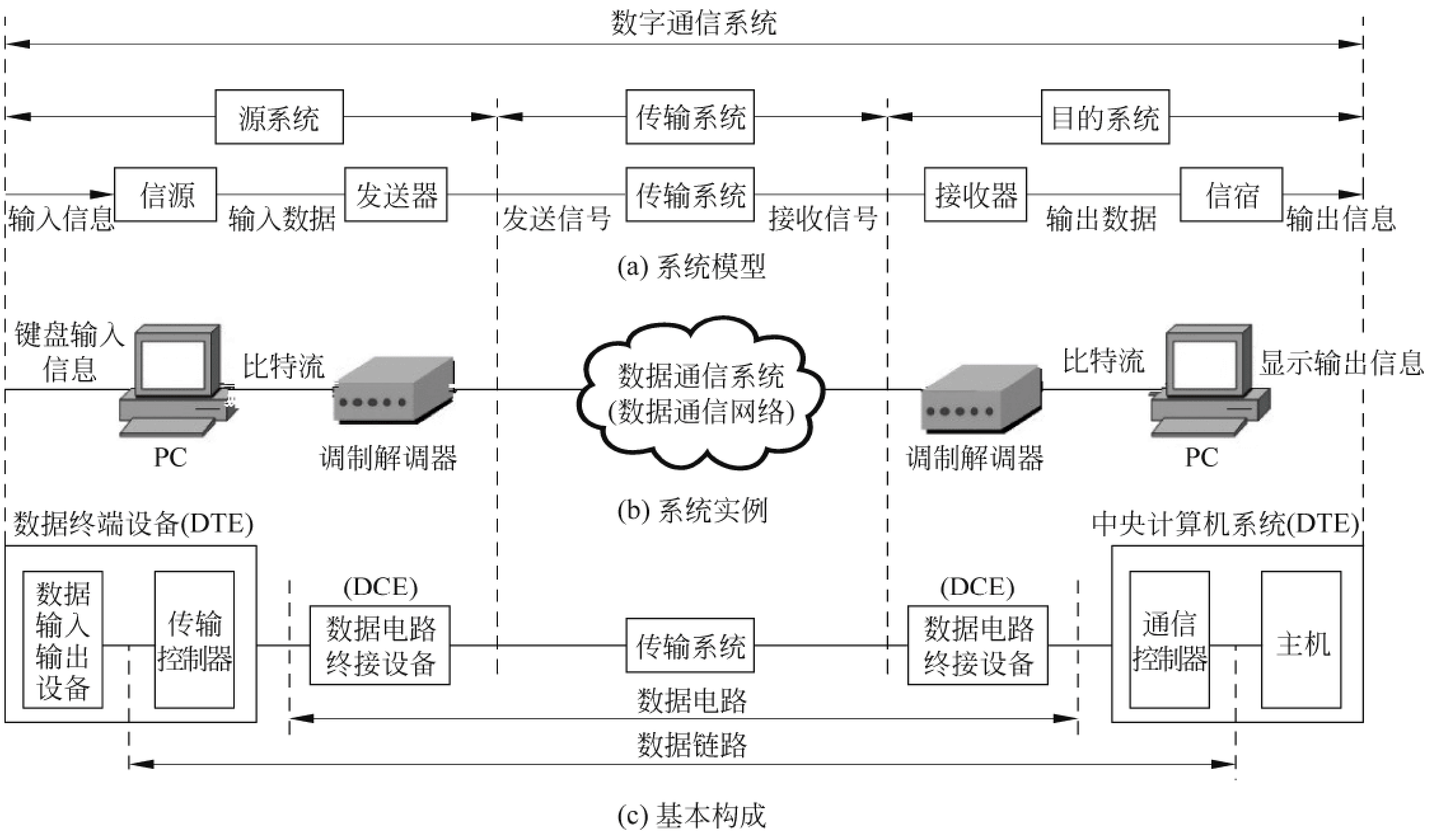


图 1.2　数据通信系统

的数据序列,减少出错的可能性;其二是把变换后的数据序列转换成能够在通信介质中传输的信号,并通过通信介质传输给其他系统。

(2) 传输系统包括通信线路和设备,负责把源系统提供的数据信号准确、及时、可靠地传输给接收系统。传输系统可以是简单的传输线路,也可以是传输网络,还可以是数据网等。

(3) 目的系统一般也由两部分组成,分别是接收器和信宿。

接收器:与发送器系统对应,主要功能有两部分。

信宿:接收来自接收器系统的数据序列,变换输出为有效的信息内容并提供给接收方用户。把接收到的数据序列还原为语音并输出给用户。

数据通信系统的实例如图 1.2(b)所示,图中的 PC 数据通信系统模型源系统的信源和信宿、调制解调器分别对应于发送器和接收器。

数据通信系统的特点:由于数据终端设备(Data Terminal Equipment,DTE)至数据电路终接设备(Data Circuit-Terminating Equipment,DCE)之间传送的是原始的二进制代码,这种二进制代码不适合长距离传送。只有将其变换成能够进行位同步,而且适合长距离传送的代码才行。为此,需要对这种二进制代码进行数字编码和模拟编码。

数字编码:如普通二进制编码(又称为非归零编码)、曼彻斯特编码、差分曼彻斯特编码等。

模拟编码:可利用调制解调器把计算机发出的数字信号转换成模拟信号。常用的调制方法有幅度调制、频率调制和相位调制。

2. 数据通信系统的基本构成

数据通信系统的基本构成如图 1.2(c)所示,由 DTE、DCE 及传输系统构成,它与典型的数据通信系统模型也有一定的对应关系。

(1) 数据终端设备由数据输入设备、传输控制器或主机、通信控制器组成,它是根据网

络实际需要采用不同设备的一个总称,目前通常将路由器或 PC 作为 DTE。

(2) 数据电路由传输系统及其两端的 DCE 组成,它的作用是为数据通信提供物理传输通道。而数据链路是从 DTE 的链路层看进去的整个链接。数据通信系统是通过数据电路将分布在远地的数据终端设备与计算机系统连接起来,实现数据传输、交换、存储和处理的系统。

(3) 数据电路终接设备是 DTE 与传输信道的接口设备。当数据信号采用不同的传输方式时,DCE 的功能有所不同。

基带传输时,DCE 对来自 DTE 的数据信号进行某些变换,使信号功率谱与信道相适应,使数据信号适合在电缆信道中传输。

频带传输时,DCE 具体是调制解调器(modem),它是调制器和解调器的结合。

当数据信号在数字信道上传输(数字数据传输)时,DCE 是数据服务单元(Data Service Unit,DSU),其功能是信号格式变换,即消除信号中的直流成分和防止长串零的编码、信号再生和定时等。图 1.3 给出了数据通信系统设备 DCE 至 DTE 常用的几种传输接口。

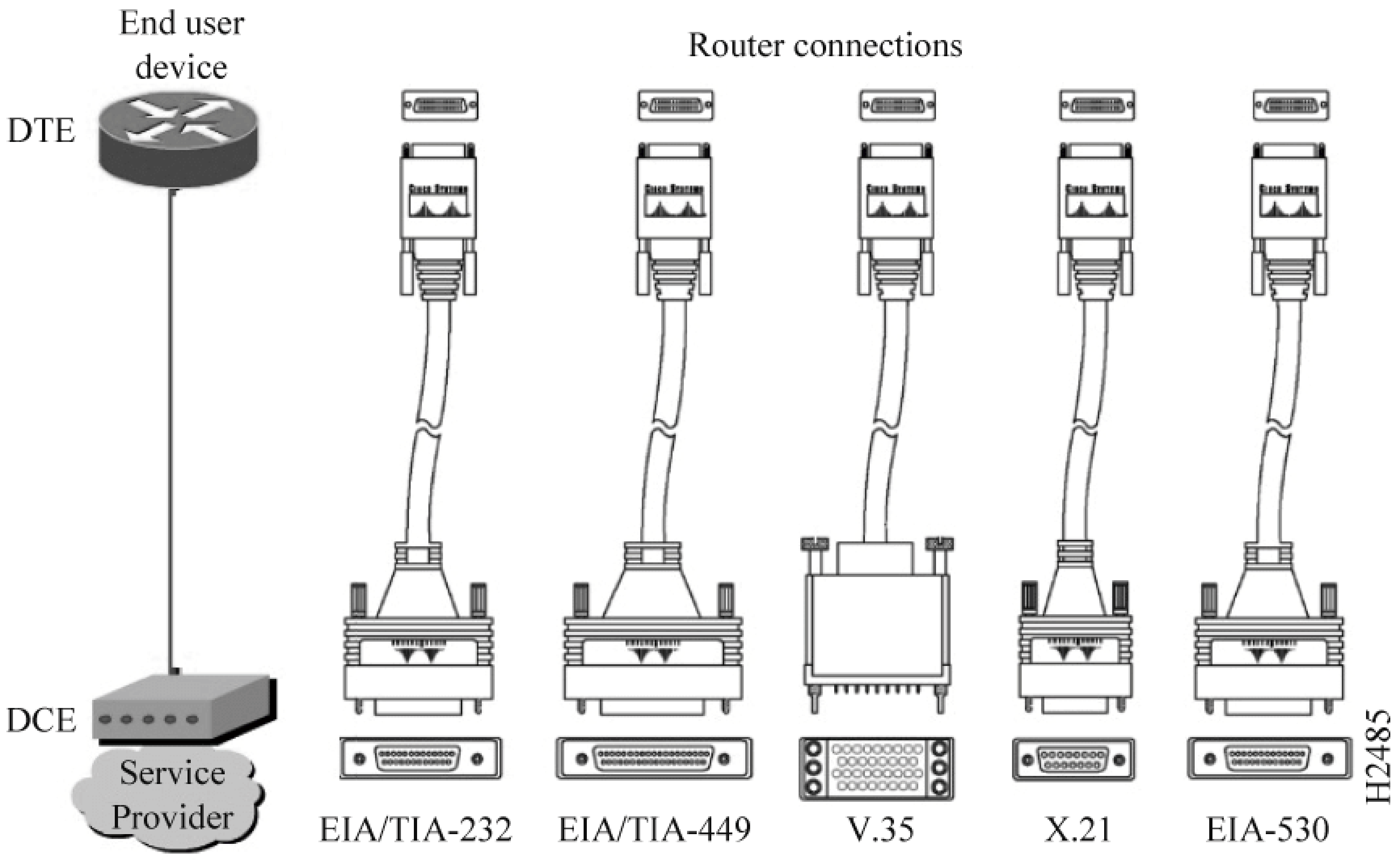


图 1.3 DCE 至 DTE 常用的传输接口

(4) 中央计算机系统由通信控制器、主机及其外围设备组成,具有处理从数据终端设备输入的数据信息,并将处理结果向相应的数据终端设备输出的功能。

通信控制器是数据电路和计算机系统的接口,控制与远程数据终端设备连接的全部通信信道,接收远端 DTE 发来的数据信号,并向远端 DTE 发送数据信号。

主机又称中央处理机,由中央处理单元(CPU)、主存储器、输入输出设备以及其他外围设备组成,其主要功能是进行数据处理。

3. 数据通信系统网络结构

为了把集散控制系统中的各个组成部分连接在一起,常常需要把整个通信系统的功能分成若干层次去实现,每一层次就是一个通信子网。通信网络的拓扑结构就是指通信网络中各个节点或站点相互连接的方法,而应用较多的是星形、环形和总线型结构。

在星形结构中,每一个节点都通过一条链路连接到一个中央节点上。任何两个节点之间的通信都要经过中央节点。一旦中央节点发生故障,整个通信系统就要瘫痪。

在环形结构中,所有的节点通过链路组成一个环形。需要发送信息的节点将信息送到环上,信息在环上只能按某一确定的方向传输。

总线型结构采用的是一种与星形和环形结构完全不同的方法,这时的通信网络仅仅是一种传输介质,它既不像星形网络中的中央节点那样具有信息交换的功能,也不像环形网络中的节点那样具有信息中继的功能,所有的站都通过相应的硬件接口直接接到总线上。由于所有的节点都共享一条公用的传输线路,所以每次只能由一个节点发送信息,信息由发送它的节点向两端扩散。总线型结构突出的特点是结构简单,便于扩充。总线型结构对总线的电气性能要求很高,对总线的长度也有一定的限制。因此,它的通信距离不可能太长。

4. 常见的数据通信系统

数据通信系统从物理层信号传输的角度看进去,一般包括模拟通信系统、数字频带传输系统、数字基带传输系统和模拟信号数字化传输系统4类。

(1) 模拟通信系统。模拟通信系统主要包含两种重要变换,即把连续消息变换成电信号和把电信号恢复成最初的连续消息的过程。经过调制后的信号通常称为已调信号。

已调信号有3个基本特性,一是携带有消息;二是适合在信道中传输;三是频谱具有带通形式,且中心频率远离零频。因而已调信号又称为频带信号。

(2) 数字频带传输通信系统。通常把有调制器/解调器的数字通信系统称为数字频带传输通信系统。数字频带通信系统的模型如图1.4所示。对于图中的调制器/解调器、加密器/解密器、编码器/解码器等环节,在具体通信系统中是否采用,取决于具体设计条件和要求。但在一个系统中,如果发送端有调制、加密、编码设备,则接收端对应有解调、解密、译码设备。

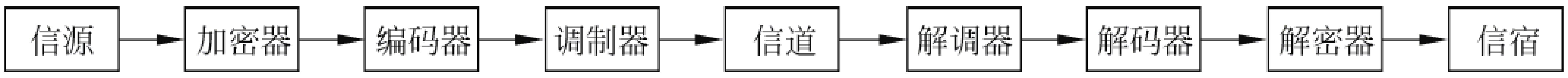


图1.4 数字频带传输通信系统

(3) 数字基带传输通信系统。与数字频带传输通信系统相对应,把没有调制器/解调器的数字通信系统称为数字基带传输通信系统,如图1.5所示。基带信号形成器包括编码器、加密器、波形变换器等,接收滤波器包括译码器、解密器等。

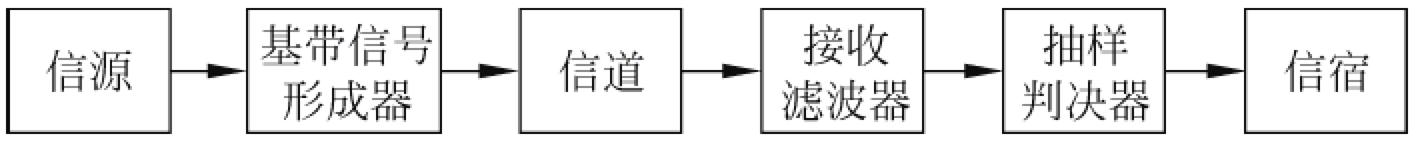


图1.5 数字基带传输通信系统

(4) 模拟信号数字化传输通信系统。要实现模拟信号在数字系统中的传输,则必须在发送端将模拟信号数字化,即进行模数(Analog/Digital,A/D)转换;在接收端需进行相反的转换,即数模(Digital/Analog,D/A)转换。

1.3.2 系统性能指标

数据通信系统可以使用数据传输速率、带宽、码元传输速率、误码率、吞吐量等性能指标,衡量数据传输的有效性和可靠性。其中有效性主要由数据传输速率、传输延迟、信道带宽、信道容量等指标来衡量;而传输系统的可靠性一般用数据传输的误码率来衡量。

1. 数据传信速率

数据传信速率(Data Transfer Rate,DTR)是每秒传输的二进制位数,又称为数据率,单位是 bps,或 kbps,Mbps,Gbps 等,其中 k(千)、M(兆)和 G(吉)分别代表 10^3、10^6 和 10^9。而通常在表示数据量时,千、兆和吉分别用 K(大写)、M 和 G 表示,分别代表 2^{10}(1024)、2^{20}(1 048 575)和 2^{30}(1 073 741 824),如 2M 口。针对单位,下面作简单概述。

b:位(bit),是计算机中最小的数据单位,每一位的状态只能是 0 或 1。

B:8 个二进制位构成 1 字节(Byte),它是存储空间的基本计量单位。

KB:这时 K 表示 1024,也就是 2 的 10 次方。如 1KB,表示 1024 字节。

MB:计量单位中的 M 是 10 的 6 次方,1MB 不正好等于 1 000 000 字节,而是 1 048 576 字节,即 1MB=2E+20B=1 048 576B。

根据进制规定,传送速度可以有两种表示方法 Bps 和 bps,但是它们有严格区别。Bps 中的 B 使用的是二进制系统中的 Byte,bps 中的 b 是十进制系统中的位元。如常说的 56K 拨号上网,1000M 局域网等,都指的是 bps 计量。当用软件下载工具时一般多以 Bps 计算,所以它们之间存在 1Byte=8bit 的换算关系,因此 56kbps 拨号的下载速度是 56kbps/8=7KBps,即每秒可以下载 7KB。

在数据存储容量计算中,一般结合公制的进制和二进制的数据计算方法来计算。如:

$$1\text{KB}=2^{10}\text{B}=1024\text{B}(\text{千字节})$$

而一些生产存储器厂家是用十进制计算。如:

$$1\text{KB}=10^{3}\text{B}=1000\text{B}(\text{千字节})$$

这就是为什么操作系统显示的容量与厂家标示的容量往往有些差异的原因。

2. 数据传送速率

指单位时间内在数据传输系统中的相应设备之间实际传送的位、字符或码组平均数,单位分别为位/秒、字符/秒或码组/秒。数据传送速率又称数据传输速率。人们一般也将数据传信速率称作数据传输速率,当然这时的单位是指位/秒。

3. 带宽

在计算机网络中,带宽(band width)用来表示数字信道所能传送的最高数据传输速率,也就是信道的最大数据传输速率,通常称为"传输带宽",常用的单位有 kbps(10^3 bps)、Mbps(10^6 bps)、Gbps(10^9 bps)、Tbps(10^{12} bps)。传输带宽与数据传输速率是有区别的,前者表示信道的最大传输速率,是信道传输数据能力的极限;而后者是实际的数据传输速率,就像高铁的最大限速与实际速度的关系一样。

4. 背板带宽

交换机的背板带宽(backplane band width),是交换机接口处理器或接口卡(包括可扩展插槽中尚未安装的板卡)和数据总线间所能吞吐的最大数据量。背板带宽标志了交换机总的数据交换能力,单位为 Gbps,也称为交换带宽。只有拥有可扩展插槽、可灵活改变端口数量的模块交换机才有这个概念,固定端口交换机是没有这个概念的,因为固定端口交换机的背板容量和交换容量大小是相等的。背板带宽决定了各板卡与交换引擎间连接带宽的最高上限。由于模块化交换机的体系结构不同,背板带宽并不能完全有效代表交换机的真正性能。背板带宽标志了交换机总的数据交换能力,一般的交换机的背板带宽从几 Gbps 到上百 Gbps 不等,一台交换机的背板带宽越高,所能处理数据的能力就越强。

5. 交换引擎转发性能

由于交换引擎(switching engine)是作为模块化交换机数据包转发的核心,所以交换引擎的转发性能(交换容量、转发能力)能够真实反映交换机的性能。对于固定端口交换机,交换引擎和网络接口模板是一体的,所以厂家提供的转发性能参数就是交换引擎的转发性能,这一指标是决定交换机性能的关键。支持第三层交换设备,厂家会分别提供第二层转发速率(bps)和第三层转发速率(pps),采用不同体系结构的模块化交换机,这两个参数的意义是不同的。但是,对于一般的局域网用户而言,只关心这两个指标就可以了,它是决定该系统性能的关键指标。

6. 波特率(码元速率)

一个数字脉冲称为一个码元,码元速率表示单位时间内通过信道传输的码元个数。在通信系统中,把承载数据的基本信号单元称为“码元”,把每秒传输的码元(符号)数称为波特率,又称为码元传输速率,单位是波特,记作 Baud。如信号码元持续时间为 $T(s)$,则码元速率

$$B = 1/T(\text{Baud}) \tag{1-1}$$

式中,B 表示波特率;T 为信号码元的时间宽度,单位为秒。

数据传输速率“位/秒”与码元的传输速率“波特”在数量上有一定的关系。例如,在计算机中,一个符号的含义分别代表逻辑“1”和逻辑“0”,所以每个符号所含的信息量刚好为 1 位,此时比特率(即每秒传送的位数)等于波特率。而采用多进制的编码方式,一个码元可以承载多位的信息,这样可以达到更高的数据传输速率。对于一个 M 进制码元,所包含的信息量为 $I=\log_2 M(\text{bit})$,所以数据传输速率的计算式如下式所示。

$$C = BI = B \times \log_2 M(\text{bps}) \tag{1-2}$$

式中,C 为比特率;B 代表波特率;M 是一个码元表示的有效状态数。

例如采用 8QAM 调制方式时,$M=8$,则信息量为 $I=\log_2 M(\text{bit})=\log_2 8(\text{bit})=3$。若波特率:$B=20\,000\text{Baud}$,则数据传输速率:$C=BI=20\,000\times 3=60\,000\text{bps}$。

波特率与比特率的关系为:

比特率=波特率×单个调制状态对应的二进制位数

需要说明:并不总是比特率大于或等于波特率,在有的情况下,比特率可能小于波特率。例如,在曼彻斯特编码中,由于一个码元应调制成两个电平,比特率就只有波特率的一半,数据传输的效率减少了一半。

7. 包转发率

包转发率标志了交换机转发数据包的能力,单位为 pps(包每秒),一般交换机的包转发率在几十 Kpps 到几百 Mpps 不等。包转发率是指交换机每秒可以转发多少个数据包,即交换机能同时转发的数据包的数量。包转发率通常在衡量三层交换机能力时用到,如在说明三层交换速率就用 pps。

8. 时延

时延(delay)是指一个报文或分组从一个网络(或一条链路)的一端传送到另一端所需的时间。网络中的时延如图 1.6 所示,一般分为以下 4 种。

(1) 发送时延(Transmission Delay,TD)。发送时延是节点在发送数据时使数据块从节点进入到传输介质所需要的时间,也就是从数据块的第一位开始发送算起,到最后一位发

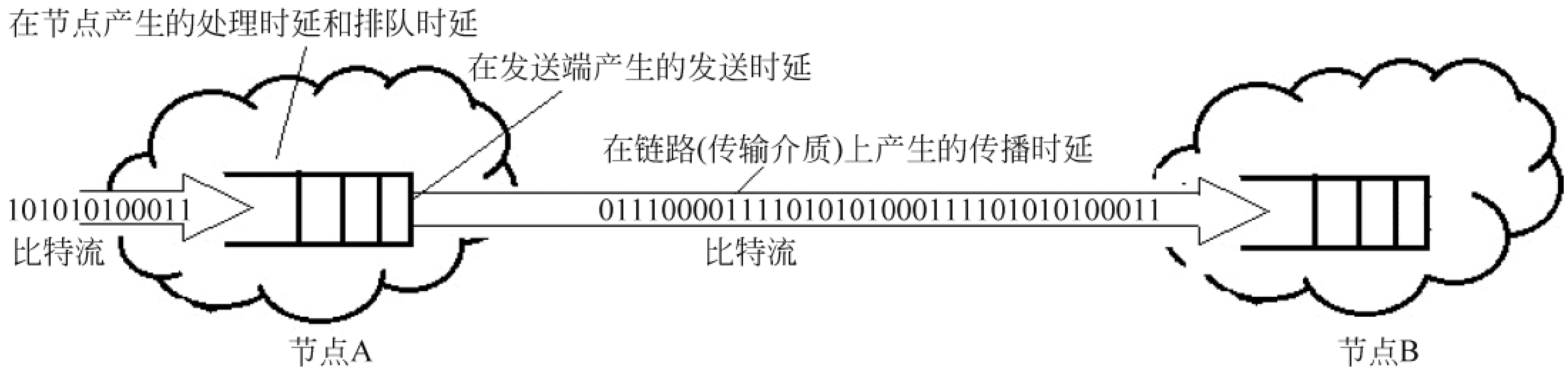


图 1.6 时延的产生

送完毕所需的时间，又称传输时延。发送时延的计算式为：

$$TD = 数据块长度(bit) \div 信道带宽(bps) \tag{1-3}$$

由此可见，对于一定的网络，发送时延与发送的数据块长度成正比，与信道带宽成反比。

(2) 传播时延(Propagation Delay，PD)。传播时延是指电磁信号或光信号，在传输介质中传播一定的距离所花费的时间，即从发送端发送数据开始，到接收端收到数据。或者理解为从接收端发送确认信号，到发送端收到确认信号，总共需要的时间，它与信号传播速度和距离有关。传播时延的计算式如下：

$$PD = 信道传播距离(m) \div 信号在信道上的传播速率(m/s) \tag{1-4}$$

信号在自由空间的传播速率是光速，即 3.0×10^5 km/s；在铜线电缆中的传播速率约为 2.3×10^5 km/s；在光纤中的传播速率约为 2.0×10^5 km/s。

【例 2】 数据长度为 10^7 位，数据发送速率为 100kbps，传播距离为 1000km，信号在介质上的传播速率为 2×10^8 m/s，试计算发送时延和传播时延。

答：发送时延：$TD=10^7/100k=100s$

传播时延：$PD=1000\times1000/(2\times10^8)=0.005s$

(3) 节点处理时延(Nodal Processing Delay，NPD)。节点在收到分组时要花费一定的时间进行处理，如对分组的封装、解封装等产生的时延。

(4) 排队时延(Queueing Delay，QD)。分组在经过网络传输时，要经过许多的路由器。分组在进入路由器后要先在输入队列中排队等待处理。在路由器确定了转发接口后，还要在输出队列中排队等待转发。这就产生了排队时延。排队时延的长短取决于网络当时的通信量，当网络的通信量很大时，还会产生分组溢出，这相当于排队时延为无穷大。

这样，数据在网络中经历的总时延(total delay)就是以上 4 种时延之和，即：

$$总时延 = TD(发送时延) + PD(传播时延) + NPD(处理时延) + QD(排队时延) \tag{1-5}$$

【例 3】 一个长度为 100MB 数据块在带宽为 1Mbps 信道上，用光纤传送到 1000km 远的目的计算机所需的时间是多少？

答：因为数据块长度，$1MB=1\,048\,576Byte=2^{20}Byte=2^{20}\times8bit$；信道带宽，$1Mbps=10^6bps$，所以根据式(1-3)可以得到：

$$发送时延=100\times2^{20}\times8/10^6=838.9s$$

又由于光纤的传播速率约为 2.0×10^5 km/s，根据式(1-4)算出光纤传送到 1000km 的

$$传播时延 = 1000(km)/200\,000(km/s) = 0.005s = 5ms$$

$$总时延 = 838.9s + 5ms = 838.905s$$

由于 ms(毫秒)级时间相对于 838.9s 来说非常小,几乎可以忽略,所以在这种情况下,总时延的数值基本上是由发送时延来决定的。

【例 4】 一个长度为 1B 数据块在带宽为 1Mbps 信道上,用光纤传送到 1000km 远的目的计算机所需的时间是多少?

答: 发送时延 $= 1 \times 8/10^6 = 8\mu s$

传播时延 $= 1000(km)/200\,000(km/s) = 0.005s = 5ms$

总时延 $= 8\mu s + 5ms = 5.008ms$

总时延中"传播时延"反而占了主要位置。即使将信道的带宽提高到 1Gbps,也只是在"发送时延"微秒级的基础上减少,因为此时的传播时延级数达到了 ms,可见在此情况下,总时延的数值是由传播时延决定的。

基于上述计算,人们经常听到的诸如"在高带宽链路上,比特流传输速度更快"和"光纤信道的传输速率高"等说法都是有问题的。因为对于高带宽,提高的仅仅是数据的发送速率,而不是比特流在链路上的传播速率。也就是说,提高链路带宽只是减少了数据的发送时延,至于传播时延还要通过信道传输速率和传输距离而定。同时光纤信道发送数据的速率可以很高,而光纤信道的传播速率实际比铜线的传播速率还略低些。

数据发送速率的单位是每秒发送多少位,是指某个点或某个接口上的发送速率;而传播速率的单位是每秒传播多少距离,是指传输线路上位的传播速率。

9. 时延带宽积

时延带宽积(Bandwidth-Delay Product,BDP)为某一链路所能容纳的位数。时延带宽积的计算式为

$$BDP = 带宽 \times 传播时延 \tag{1-6}$$

时延带宽积表示的是以位为单位的链路长度。如设某段链路的传播时延为 10ms,带宽为 10Mbps,则时延带宽积$=10\times10^{-3}\times10\times10^{6}=1\times10^{5}$位。这表明,若发送端连续发送数据,则当发送的第 1 位即将到达接收端时,发送端就已经在链路上发送了 10 万位。

10. 利用率

利用率(utilization)有信道利用率和网络利用率两种。信道利用率是指某信道有百分之几的时间是被利用的(有数据通过),完全空闲的信道的利用率是零。网络利用率则是指全网络的信道利用率的加权平均值。

信道利用率是一把双刃剑,互联网服务提供商(Internet Service Provider,ISP)希望高一些好,它能在一定程度上体现网络设备的利用率和用户数的增长率。而作为用户来说并非越高越好,因为根据排队论,当某信道的利用率增大时,该信道引起的时延也就迅速增加,网络的服务质量就会下降。网络的信道利用率高低,就类似于高速公路上车流量大小的情况,当网络的通信量很少时,网络产生的时延并不大;但在网络通信量较大的情况下,由于分组在网络节点(路由器或交换机)进行处理时需要排队等候,因此网络引起的时延就会增大。

在适当的假定条件下,如果令 D 表示网络当前的时延,则:

$$D = D_0 \div (1 - U) \tag{1-7}$$

式中,D_0 表示网络空闲时的时延;U 是网络的利用率,数值在 0~1 之间。当网络的利用率达到其容量 1/2 时,时延就要加倍,而当网络的利用率接近最大值 1 时,网络的时延就趋于无穷大。因此,应该有这样的概念:信道或网络利用率过高会产生非常大的时延。因此一

些拥有较大主干网的 ISP 通常控制它们的信道利用率不超过 50%,如果超过了就要准备扩容,增大线路的带宽。

11. 吞吐量

吞吐量(throughput)表示在单位时间内通过某个网络(或信道、接口)的数据量。吞吐量受网络带宽或网络额定速率的限制,是经常用于测量网络通过数据量的一项指标。例如,对于一个 100Mbps 的以太网,其吞吐量的绝对上限值就是 100Mbps,而实际的吞吐量可能只有 60Mbps。要注意的是,有时吞吐量还可以每秒传送的字节数或帧数来表示。

12. 每秒的输入输出量

每秒读写次数(Input/Output Per Second,IOPS),是指系统在单位时间内能处理的最大的 I/O 频度,是指单位时间内系统能处理的 I/O 请求数量,I/O 请求通常为读或写数据操作请求。对于大量顺序读写,如视频点播(Video On Demand,VOD),则更关注吞吐量指标。

13. 误码率

误码率(Bit Error Ratio,BER)是指二进制数字位传输时出错的概率,是衡量数据在规定时间内传输可靠性的指标。根据概率统计理论,误码率为

$$\mathrm{BER} = N_e \div N \times 100\% \tag{1-8}$$

式中,N_e 为被传错的码元数;N 为所传送的码元总数。如果实际传输的不是二进制码元,需折合成二进制码元计算。

14. 差错率

数据信号在传输过程中,噪声干扰和信号畸变达到一定程度时就可能导致接收的差错。衡量数据传输质量的最终指标是差错率。

在数据传输中,一般采用误码(位)率、误字符率、误码组率,它们分别定义如下:

误码(位)率 = 接收出现差错的位数 / 总的发送位数

误字符率 = 接收出现差错的字符数 / 总的发送字符数

误码组率 = 接收出现差错的码组数 / 总的发送码组数

差错率是一个统计平均值,因此在测量或统计时,总的位(字符、码组)数应达到一定的数量,否则得出的结果将失去意义。

15. 频带利用率 η

数据信号的传输需要一定的频带。数据传输系统占用的频带越宽,传输数据信息的能力越大。即使两个数据传输系统的传信速率相同,但它们的通信效率也可能不同,这就需要看传输相同信息所占的频带宽度,所以频带利用率定义为单位频带内的码元速率,即每赫兹的波特数:

$$\eta = \text{系统的码元速率} / \text{系统的频带宽度} \tag{1-9}$$

当然,衡量数据传输系统有效性的指标还可以是单位频带内的传信速率,即每赫兹每秒的位数(bit/(s · Hz))。

1.4 数据交换技术

从通信资源的分配角度来看,“交换”就是按照某种方式动态地分配传输线路的资源。在计算机网络及通信系统中常谈到的交换技术有电路交换、分组交换、ATM 交换及软交换

等。从最初适应语音通信的电路交换,到适应数据通信的分组交换,又发展到适应于宽带综合数据业务的ATM。为了适应移动互联网及NGN(下一代网络)的发展,又推出软交换、IMS、MPLS及ASON等新技术。各种交换方式发展关系如图1.7所示。

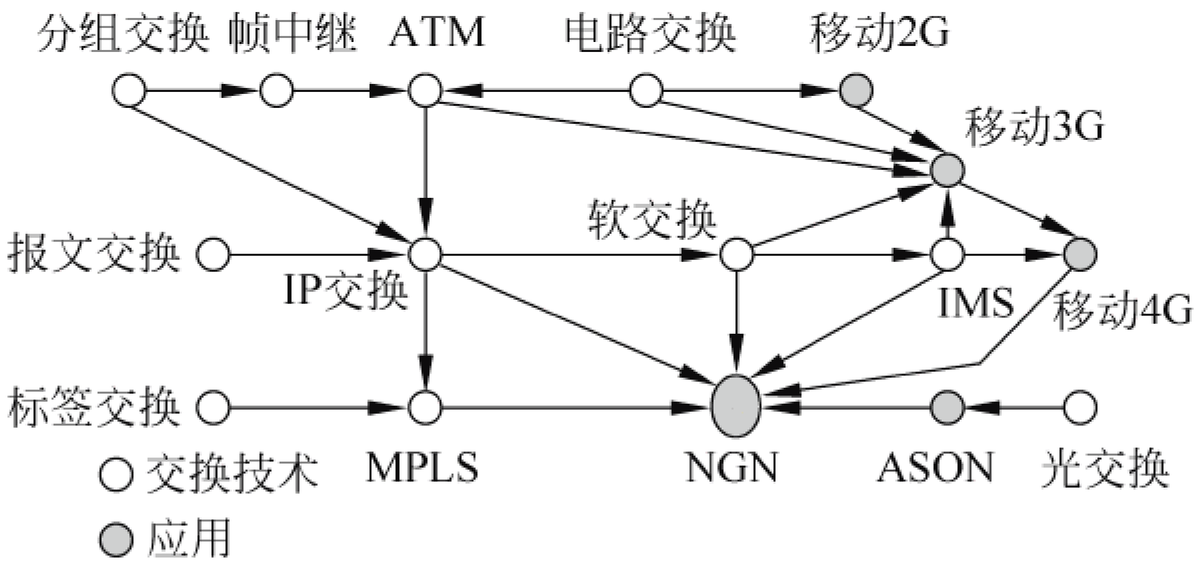


图1.7 各种交换方式的发展关系及应用

1. 电路交换

电路交换(Circuit Switching,CS)是一种直接的交换方式,在通信之前要在通信双方之间建立一条被双方独占的物理通道,这条通道是由节点内部电路对节点间传输路径通过适当选择、连接而形成的。电路交换提供给用户的是"透明通道",即交换网对用户信息的编码方法、信息格式以及传输控制程序等都不加以限制,但对通信双方来说,必须做到双方的收发速度、编码方法、信息格式以及传输控制等完全一致才能完成通信。

电路交换的特点:通信用户间必须建立专用的物理连接通路,这个连接过程只要不释放,物理连接就永远保持,在通信结束之后释放链路。在整个通信进行的过程中,通信信道由参与通信的用户独享,即使某个时刻没有信息在信道上传递,其他用户也不能使用此信道。采用这种交换方式,可以保证用户的通信带宽,时延较短;但线路的利用率不高。现在广泛使用的电话通信网络中使用的就是电路交换方式,如PSTN交换机以及2G、3G的MSC等,都采用的是电路交换技术。

2. 报文交换

报文交换(Message Switching,MS)是以报文为数据交换的单位,报文携带有目标地址、源地址等信息,在交换节点采用存储转发的传输方式。报文交换采用存储-转发的传输方式,不需要为通信双方预先建立一条专用的通信线路,用户可随时发送报文,是一种无连接服务。报文交换方式适用于实现不同速率、不同协议的终端间或点对多点的传输,是以报文为单位进行存储转发的数据通信。由于报文交换传输时延大、占用存储空间大,因而不适用要求网络时延较小的数据通信。数据通信中有一些交换机及协议采用报文交换技术。

3. 分组交换

分组交换(Packet Switching,PS)采用存储转发方式,将用户要传送的报文分成若干组,以减少存储时间。分组是指包含用户数据和协议头的块,每个分组通过网络交换机或路由器被传送到正确目的地。一条信息可能被划分为多个分组,每个分组在网路中独立传输,并且可能沿不同路由到达目的地。一旦属于同一条信息的所有分组都到达了目的地,就可以将它们重装,形成原始信息,传递给上层用户。分组交换具体分为数据报传输分组交换和虚电路传输分组交换。

数据报传输分组交换:交换设备将进入网络的任一分组都作为单独的小报文来处理,

而不去理会它究竟是属于哪个报文的分组，人们将这些作为基本传输单位的“小报文”称为数据报，其交换原理类似于报文交换，而这种交换多发生在网络层，有些操作都交给上一层来完成。如 IP 协议就属于这种交换方式，但它对这些所谓的“小报文”还可以再进行切片。

虚电路传输分组交换：是指两个用户的终端设备在开始互相收发数据之前需要通过通信网络建立逻辑上的连接，这种连接直至用户不需要收发数据时才被清除。X.25 协议就属于这种分组交换，其主要特点是所有分组都必须沿着事先建立的虚电路传输，是一种面向连接服务。这种方法对于数据量较大的通信来说具有传输率高、分组传输时延小和不容易产生分组丢失等优点，但它存在对网络依赖性大的缺点。

4. 帧交换

发生在数据链路层的交换，如互联网设备中的网桥、以太网交换机等都属于帧交换，人们通常称它为 MAC(媒体接入控制)交换。最有代表性的帧交换(FS)是在改进 X.25 协议后发展而成的帧中继(FR)，它只有下面两层，没有第三层，所以加快了处理速度。通常在第三层上传输的数据单元称为分组，在第二层上传输的数据单元称为帧(frame)，FR 在数据链上以简化的方式来传输和交换数据单元。

5. ATM

ATM(异步传送模式)是 ITU-T 确定用于宽带综合业务数字网(B-ISDN)的复用、传输和交换模式技术。ATM 在综合了电路交换和分组交换优点的同时，克服了电路交换方式中网络资源利用率低、分组交换方式信息时延大和抖动的缺点，提高了网络的效率。ATM 的传输过程分为建立连接、数据传输和连接终止 3 个阶段。ATM 提供高速、高服务质量的信息交换，灵活的带宽分配及适应从很低速率到很高速率的带宽业务。

6. 移动通信交换

移动通信(PLMN，公共陆地移动通信网络)的核心网：2G、3G(R99)采用电路交换，3G(R4)采用软交换及 ATM，3G(R5 以后版本)以及 4G(R9 以后版本)采用 ISM(IP 多媒体子系统)。无线接入通信系统主要采用频分复用多址(FDMA)、时分复用多址(TDMA)和码分复用多址(CDMA)、正交频分复用多址(OFDMA)和非正交复用多址(NDMA)等技术。

7. IP 交换

一直以来，业界关于建设宽带传输的核心交换技术存在两个发展方向：一个是计算机界推崇的 IP 技术，另一个是电信界倡导的 ATM 技术。目前将 ATM 高速交换技术作为第二层与第三层的 IP 路由技术的优点结合起来，形成了适用于 IP 技术的各种路由交换设备，如 ATM 网上运行 IP(IPOA)、局域网仿真(LANE)以及 ATM 上的多协议(MPOA)等。

8. MPLS

在标签交换的基础上发展起来的 MPLS(多协议标签交换)，既具有 ATM 的高速性能，又具有 IP 的灵活性和可扩充性，可以在同一网络中同时提供 ATM 和 IP 业务。利用 ATM 传送 IP 是目前公用骨干网上最适用的技术方案之一。MPLS 已成为业界普遍看好的下一代 IP 骨干网技术。

9. 软交换

软交换(SS)概念是较早引入到下一代交换网络(NGN)的一项技术，在传输为分组网的前提下，能一定程度上实现业务、控制和承载的分离。软交换技术是一种分布的软件系统，可以基于各种不同技术、协议和设备，在网络环境之间提供无缝的互操作功能。软交换技术

通过相应的协议控制或通信规程支持 IP PBX 和 IP 电话，同时它还具有网关处理能力。软交换设备是分组网络的核心设备，它独立于网络，主要完成呼叫控制、资源分配、协议处理等功能，可以提供包括现在电路交换机所提供的全部业务和其他新的业务。

10. IMS

IP 多媒体子系统体系(IMS)结构设计利用了软交换技术，实现了业务与控制相分离、呼叫控制与媒体传输相分离。IMS 虽然是 3GPP 为了移动用户接入多媒体服务而开发的系统，但由于它全面融合了 IP 域的技术，并在开发阶段就和其他组织进行密切合作，使得 IMS 实际已经不仅仅局限于只为移动用户进行服务。

11. 光交换

经过多年的广泛研究，WDM(波分复用)技术在光网络中日趋成熟，全光交叉连接设备(OXC)和全光分插复用设备(OADM)已经得到了设计应用，光信号可以根据其波长直接在光网络中确定路由，而不需要进行光—电—光的转换。ASON(自动交换光网络)将 IP 传输网的智能性和 WDM 光网络的宽带有机地结合在了一起。现在商用单波光纤的传输容量可以达到 10Gbps 以上，如果采用光复用技术，一根光纤的传送容量至少可以达到 2000Gbps 以上，这就为人们研究全光交换网络技术带来了极大的诱惑。目前用到的技术有波分光交换、时分光交换、自由空间光交换及混合光交换等。

【例 5】 交换节点组网由图 1.8 所示，假设两个用户 A、B 之间的传输通路，经过交换节点 1 和交换节点 2，中间由 3 个网段组成，每段的传输延迟为 10^{-3}s，电路交换或分组交换(虚电路)时的呼叫建立时间为 0.1s，在这样的线路上传输 3200 位的一个报文，每个分组为 1024 位，报头或分组头的开销为 16 位，线路的数据速率是 9600bps，假设中间节点的处理时延为零，试分别计算在下列各种交换方式下，两个用户之间的延迟时间：①电路交换；②报文交换；③分组交换(虚电路)；④IP 交换(数据包)。

图 1.8 交换节点组网

答：

(1) 电路交换时，要考虑呼叫建立时间 0.1s、三段电路的时延 3×10^{-3}s，以及一个报文的发送时延 3200/9600s，所以两个用户之间的延迟时间为

$$T = 0.1 + 3200/9600 + 3 \times 10^{-3} = 0.436\text{s}$$

(2) 报文交换时，是一种无连接的服务，没有呼叫建立时间，但要考虑增加报头开销 16 位，则终端 A、节点 1、节点 2 共 3 个设备的报文发送时延为 $3\times(3200+16)/9600$s，三段电路的时延 3×10^{-3}s，所以两个用户之间的延迟时间为

$$T = 3 \times (3200 + 16)/9600 + 3 \times 10^{-3} = 1.008\text{s}$$

(3) 对于分组交换，也就是虚电路链接时，是面向连接服务，呼叫建立时间为 0.1s，不需要考虑增加报头开销，则终端 A 的报文发送时延为$(3200+16)/9600$s，节点 1、节点 2 共形成的分组发送时延为 $2\times(1024/9600)$s，三段电路的时延 3×10^{-3}s，所以两个用户之间的延迟时间为

$$T = 0.1 + 3200/9600 + 3 \times 10^{-3} + 2 \times (1024/9600) = 0.65\text{s}$$

(4) IP 交换时,也就是数据包的传送,由于每个分组为 1024 位,终端 A 需要将 3200 位的一个报文分成 4 个分组,则每个分组需要加 16 位的分组头,则 A 的发送时延为(3200+16×4)/9600s,节点 1、节点 2 共形成的分组发送时延为 2×(1024+16)/9600)s,三段电路的时延 3×10^{-3} s,所以两个用户之间的延迟时间为

$$T=(3200+16\times4)/9600+2\times(1024+16)/9600+3\times10^{-3}=0.56\text{s}$$

1.5 数据传输网技术

传输网就是能满足各种业务和信号传输的统一平台。随着光纤到户的逐步发展以及宽带主干网的快速推进,以光纤为介质的传输网 SDH、WDM 和 PTN 就显得尤为重要。

1.5.1 SDH

同步数字体系(Synchronous Digital Hierarchy,SDH)构成了世界性的、统一的 NNI 的基础,因为 SDH 除了支持基于电路交换的同步传输模式(STM)外,还可支持基于分组交换的异步转移模式(ATM)。支持 IP 的多业务数据平台(MSTP)就是通过 SDH 实现的,而新出厂的 SDH 完全具备了多业务接入及传输,下面只介绍传统 SDH。

1. SDH 组成

SDH 信号的基本同步传送模块的速率为 155.520Mbps(STM-1),更高速率等级的同步数字系列信号是 STM-N(N=4、16、64),下面列出了 N=1、4、16、64 时的线路码速。

第 1 级为 STM-1,线路码速为 155.520Mbps。

第 2 级为 STM-4,线路码速为 622.080Mbps。

第 3 级为 STM-16,线路码速为 2488.320Mbps。

第 4 级为 STM-64,线路码速为 9953.280Mbps。

SDH 的基本网络单元有同步光缆线路系统、同步复用器(SM)、终端复用器(TM)、数字交叉连接设备(DXC)、光中继器(REG)、分插复用器(ADM)和同步数字交叉连接设备(SDXC)等。如图 1.9 所示是常用的 SDH 网络单元。

(1) 终端复用器(TM)。如图 1.9(a)所示,是双接口器件,用于网络终端站。将低速支路信号复用进 STM-N 帧上的任意位置,或完成相反的变换。

(2) 再生中继器(REG)。如图 1.9(b)所示,REG 有两种:一种是纯光的再生中继器,主要进行光功率放大以延长光传输距离;另一种是电再生中继器,属双接口器件,只有两个线路接口。它通过光/电转换、电信号抽样判决再生整形、电/光转换,以达到消除线路噪声积累的目的,保证线路上传送信号波形的完好。

(3) SDH 数字交叉连接设备(DXC)。如图 1.9(c)所示,适用于 SDH 的 DXC,称为 SDXC。SDXC 是能在接口端间提供可控 VC 的透明连接和再连接的设备,其接口速率既可以是 SDH 速率,也可以是 PDH 速率。此外,它具有一定的控制、管理功能。SDXC 的输入/输出接口与传输系统相连。

DXC 的核心部分是交叉连接功能,参与交叉连接的速率一般等于接入速率。交叉连接速率与接入速率之间的转换需要由复用和解复用功能来完成。例如,将若干个 2Mbps 信号复用至 155Mbps 中或从 155Mbps、140Mbps 中解复用出 2Mbps 信号;分离本地交换业务

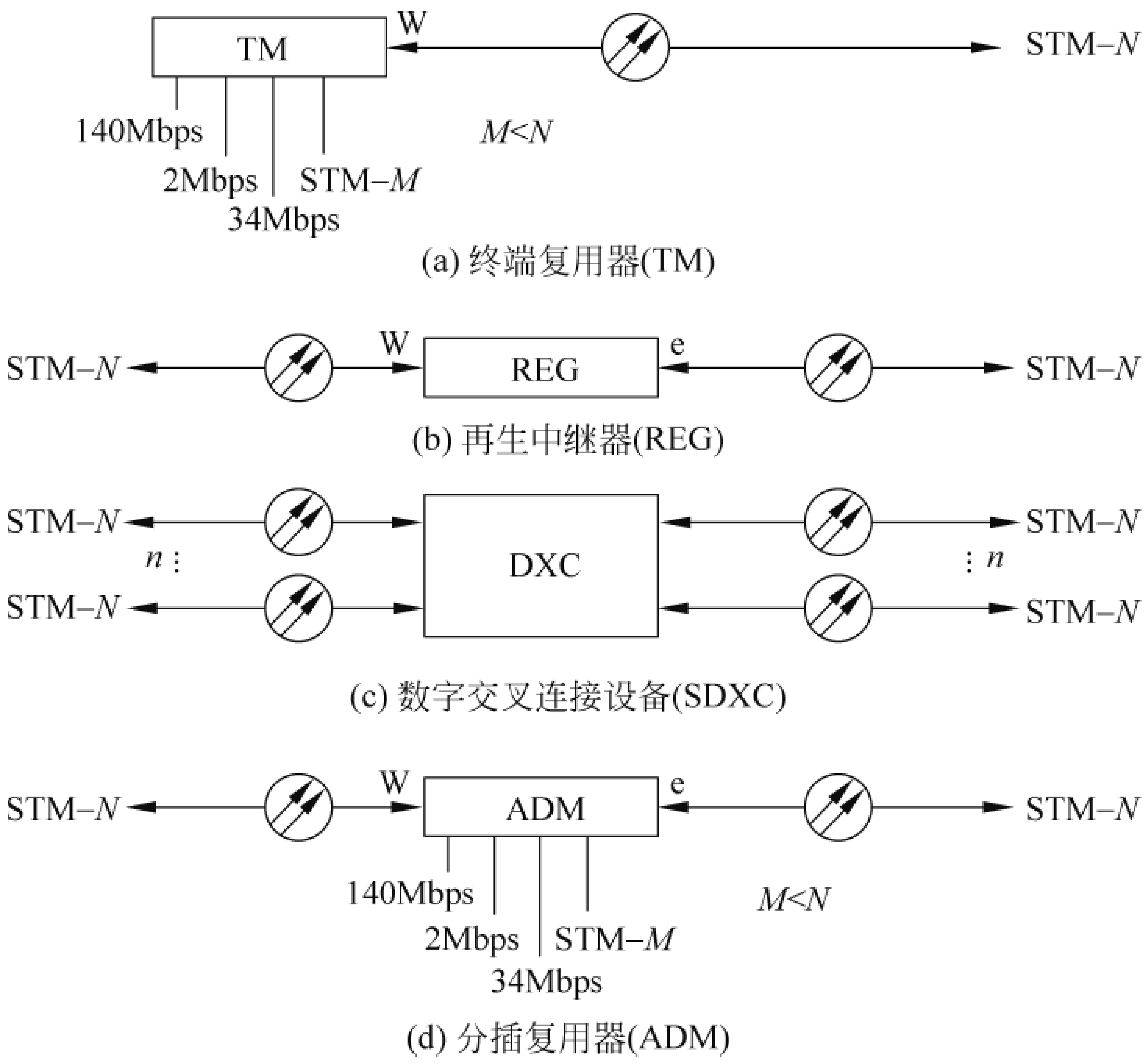


图 1.9　SDH 网络单元

和非本地交换业务，为非本地交换业务迅速提供可用路由数字交叉连接设备，是一种环间互连设备。

DXC 通常用 DXC X/Y 表示，X 表示接入接口数字流的最高等级，Y 表示参与交叉连接的最低级别。X、Y 具体取值时，数字 0 表示 64Kbps 的速率，数字 1、2、3、4 分别表示 PDH 中一至四次群的速率，其中 4 代表 SDH 中的 STM-1 速率，数字 5 和 6 分别代表 SDH 体制中的 STM-4 和 STM-16 速率。例如，DXC 4/0 表示接入接口最高速率为 STM-1 信号，而交叉连接的最低级别速率则为 64Kbps。

(4) 分插复用器(ADM)。如图 1.9(d)所示，用于 SDH 传输网络的转接站点处，它是一个三接口的器件。ADM 有两个线路接口和一个支路接口。两个线路接口各接一侧的光缆(每侧收/发共两根光纤)。ADM 的作用是将低速支路信号交叉复用进两侧线路(即上电路)，或从线路接口收的线路信号中拆分出低速支路信号(即下电路)。

采用 ADM 可以在各网络层之间提供网间连接，灵活分配不同带宽和各种业务支路接口，并且可以作为小容量交叉连接设备使用。利用 ADM 还可以构成自愈环，提供有效的线路和通道保护。

(5) 光纤连接器及光模块。连接器件主要指连接光纤适配器、光纤终端盒、路由器、交换机等网络设备的接口与线缆的组合件，通过这些部件来构成布线系统中各种子系统，组成易于实施，也能随需求变化而进行互换或升级。常见光接口连接器如图 1.10 所示，其光纤接口与接头如下。

SC/PC：方型光纤接头/微凸球面研磨抛光，其形状如图 1.10(a)所示。

ST：卡接式圆形光纤接头，其形状如图 1.10(b)所示。

FC/PC：圆形光纤接头/微凸球面研磨抛光，其形状如图 1.10(c)所示。

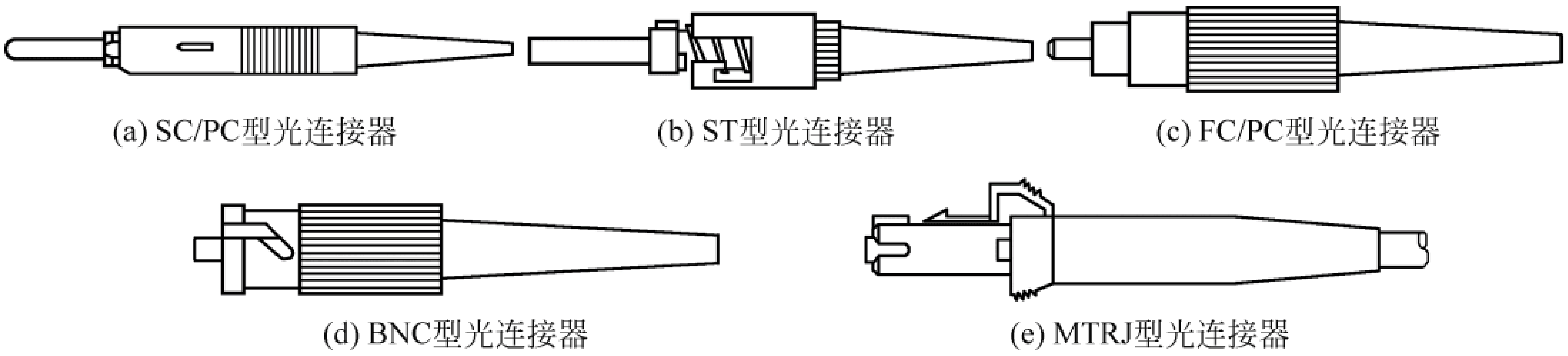


图 1.10 常见光接口连接器

BNC：多模光纤连接器，其形状如图 1.10(d)所示。

MTRJ：方形、一头双纤、收发一体常见光纤接头，其形状如图 1.10(e)所示。

FC/APC：圆形光纤接头，呈 8°角并作微凸球面研磨抛光。

一般来说，不同类型的光纤接头和相配套的光接口连接器是配套使用的。另外，以太网交换机常配的光模块如下。

SFP(Small Form-factor Pluggable transceiver)：小封装可插拔收发器。

GBIC(Gigabit Interface Converter)：千兆以太网接口转换器。

XFP(10 Gigabit small Form-factor Pluggable transceiver)：以太网接口小封装可插拔收发器。

XENPAK(10 Gigabit EtherNet transceiver PAcKage)：万兆以太网接口收发器集合封装。

2. SDH 组网

随着光纤传输容量的增大，传输网络的可靠性、可用性和对线路故障的应变能力至关重要，如图 1.11 所示给出了 SDH 传输网的各种结构，并在 SDH 传送网中采取了一系列保护机制，如图 1.12 所示为链状网和环形网采用简单的主/被用方式(即 1+1)保护措施。

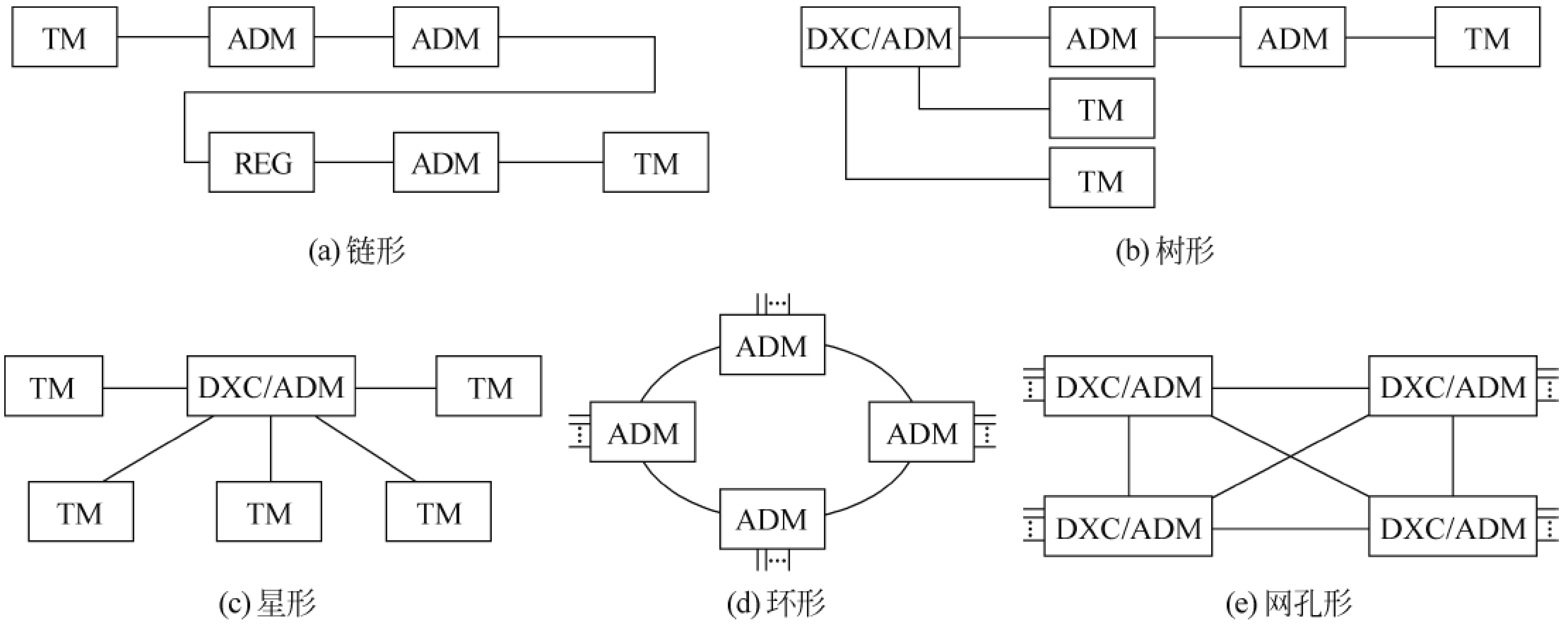


图 1.11 SDH 传输网的各种结构

首先，SDH 网络拓扑的选择应综合考虑网络的生存性作为一般性原则：星形和环形适用于用户网；线形和环形适用于中继网；树形、网孔形以及两者的结合适用于长途网。其次是倒换环的选择，通道倒换环的业务量保护是以通道为基础的，复用段倒换环的业务量是以复用段为基础的。前者根据离开环个别通道的信号质量优劣决定是否倒换，后者根据每

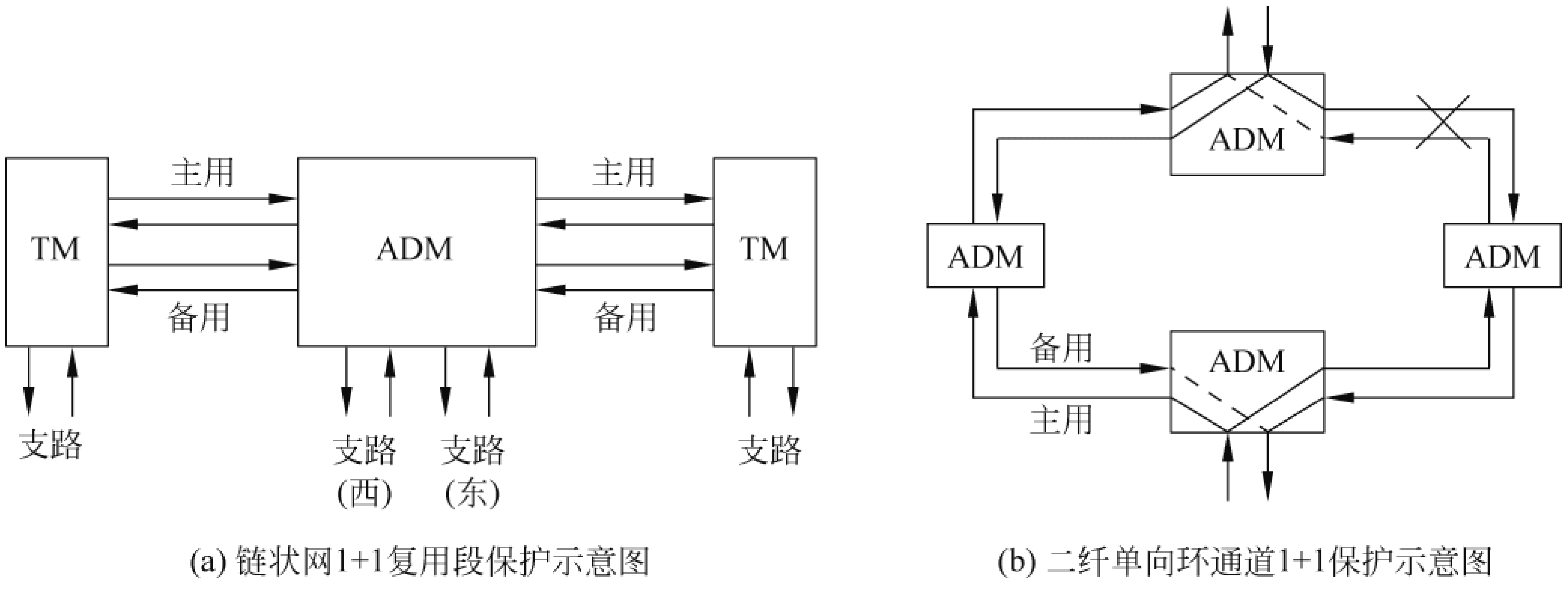


图 1.12　SDH 网络 1+1 保护

一对节点间的复用段信号质量的优劣来决定是否倒换。

1.5.2　WDM

在光波分复用(Wavelength Division Multiplexing,WDM)平台上组网克服了再生段、复用段等距离因素的限制,组网灵活,接口丰富。WDM 有两种交换方式:光路交换(circuit-switching)和光分组交换(packet-switching),由此形成了两种光波分复用网络的网络形式,以及光路交换 WDM 网(Circuit-switching WDM Network)和分组交换 WDM 网(Packet-switching WDM Network)。光交换的全光 WDM 网络有两种主要的形式:一种是广波选择网络,也就是常说的星形结构的网络;另一种是波长寻径网络。

1. WDM 节点技术

WDM 网络节点技术主要包括光交叉连接节点(OXC)、光分插复用节点(OADM)和混合节点(同时具有 OXC 和 OADM 功能的节点)等。

OXC 类似于 SDH 网络中的数字交叉连接(DXC),只不过在光域上实现,无须进行光电/电光转换和电信号处理。OXC 节点又分为静态 OXC 节点和动态 OXC 节点。在静态 OXC 节点中,不同光路信号的物理连接状态是固定的,其技术实现的难度比较小。在动态 OXC 节点中,不同光路信号的物理连接状态可以根据需要进行实时改变。

OADM 似于 SDH 网络中的分插复用器(ADM),但是直接以光波信号为操作对象,利用光波技术在光域上实现传统的电 SDH 分插复用在时域内完成的功能。

2. WDM 系统

1) WDM 常见光纤

G.652 光纤:即常规光纤(SMF),如果使用色散调制技术(如 DCF 法),则可有效地抵消光纤的色散,实现超过几千千米的长距离安全光传输。

G.653 光纤:又称色散位移光纤(DSF),是最佳的应用于单波长远距离传输的光纤。

G.655 光纤:又称非零色散位移光纤(NZDSF),它不太适合于 WDM 系统,而后来开发的 G.653 光纤更适合于 WDM 系统的应用。

2) WDM 系统

WDM 系统可分为集成式 WDM 系统和开放式 WDM 系统两大类。

集成式 WDM 系统是指 SDH 中继必须具有满足 G.692 的光接口,包括标准的光波长

和满足长距离传输的光源。这两项指标都是当前 SDH 系统(G.957 接口)不要求的,即需要把标准的光波长和长色散受限距离的光源集成在 SDH 系统中。对于集成式 WDM 系统中的 TM、ADM 和 REG 设备,都应具有符合 WDM 系统要求的光接口,以满足传输系统的要求,整个系统构造没有增加多余设备,如图 1.13(a)所示。

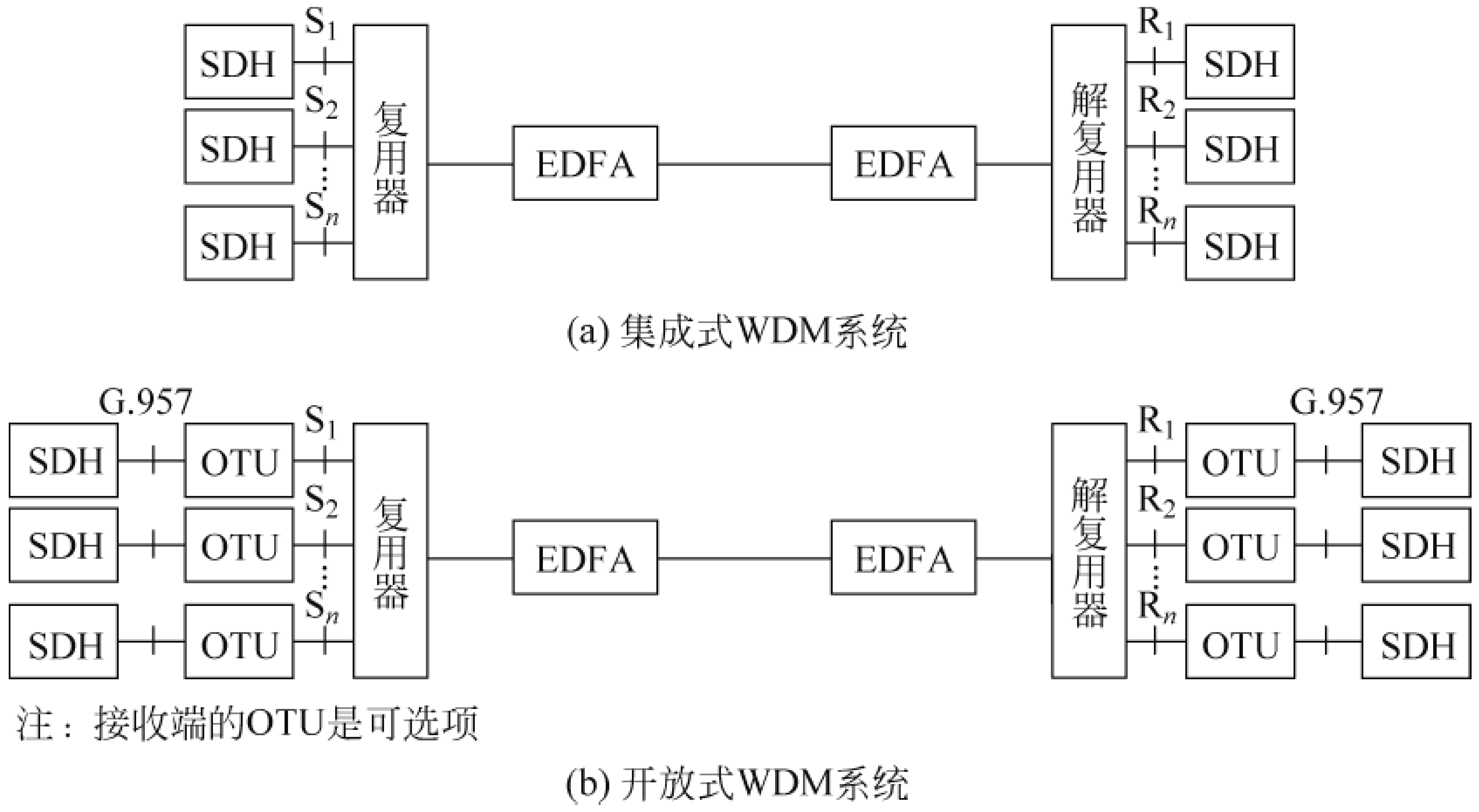


图 1.13　两类 WDM 系统示意图

开放式 WDM 系统是指发送端设备有光波长转发器(OTU),它的作用是在不改变光信号数据格式的情况下,把光波长按照一定的要求重新转换,以满足 WDM 系统的设计要求,如图 1.13(b)所示。

OTU 对输入端的信号波长没有特殊要求,接纳过去的 SDH 系统,实现不同厂家 SDH 系统工作在一个 WDM 系统内。OTU 输出端满足 G.692 的光接口,即标准的光波长和满足长距离传输的光源。具有 OTU 的 WDM 系统不再要求 SDH 系统具有 G.692 接口,可继续使用符合 G.957 接口的 SDH 设备。OTU 可看作是 WDM 系统的网元,通过 WDM 的网元管理系统进行配置和管理。

在实际建网中,可以根据要求选取系统。例如,在多厂商 SDH 系统的环境中,可以选择开放式系统,而新建干线和 SDH 系统较少的地区,可以选择集成式系统,以降低成本。

1.5.3　PTN

1. PTN 的网络分层

PTN 的网络分层结构如图 1.14 所示,主要由 3 层网络组成,它们是传输介质层、虚通路(VP)层和虚通道(VC)层。对于采用 MPLS-TP 技术的 PTN 而言,VC 层即为 PW(伪线)层;VP 层即为标签交换路径 LSP 层。传输介质层可采用以太网、SDH 等传输技术。客户业务层在 PTN 网络的最上层,可以是绑定的多个客户或是基于接口的客户。

2. PTN 网元分类

PTN 网元可分为网络边缘节点(PE)和核心节点(P)两种类型,如图 1.15 所示。用户边缘设备(CE)是进出 PTN 网络业务层的源、宿节点,在 PTN 网络的两端成对出现。P 节点是在 PTN 网络内部进行 VP 隧道转发的网元。PE 和 P 描述的是对客户业务、VC(PW)、VP(LSP)的逻辑处理功能。对于一个指定的分组网络传送业务,PE 或 P 的功能只能被一

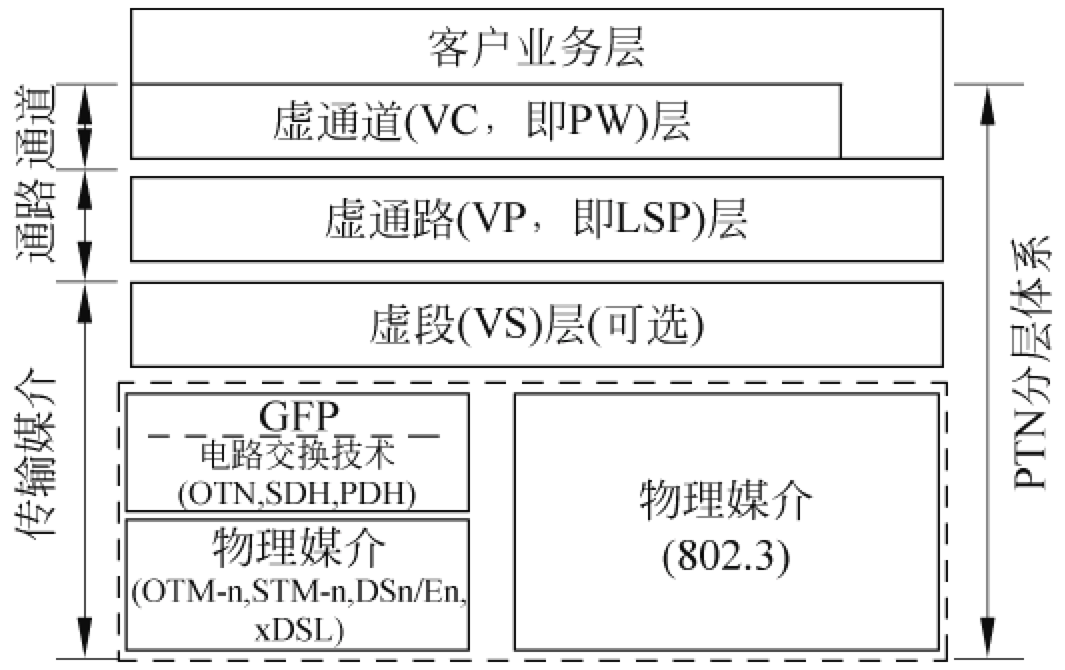


图 1.14　PTN 的网络分层结构

个特定的 PTN 网元所承担。但从任何一个 PTN 网元上来看，可以同时承载多条分组传送网业务，因而该 PTN 网元可以是 PE，也可以是 P。

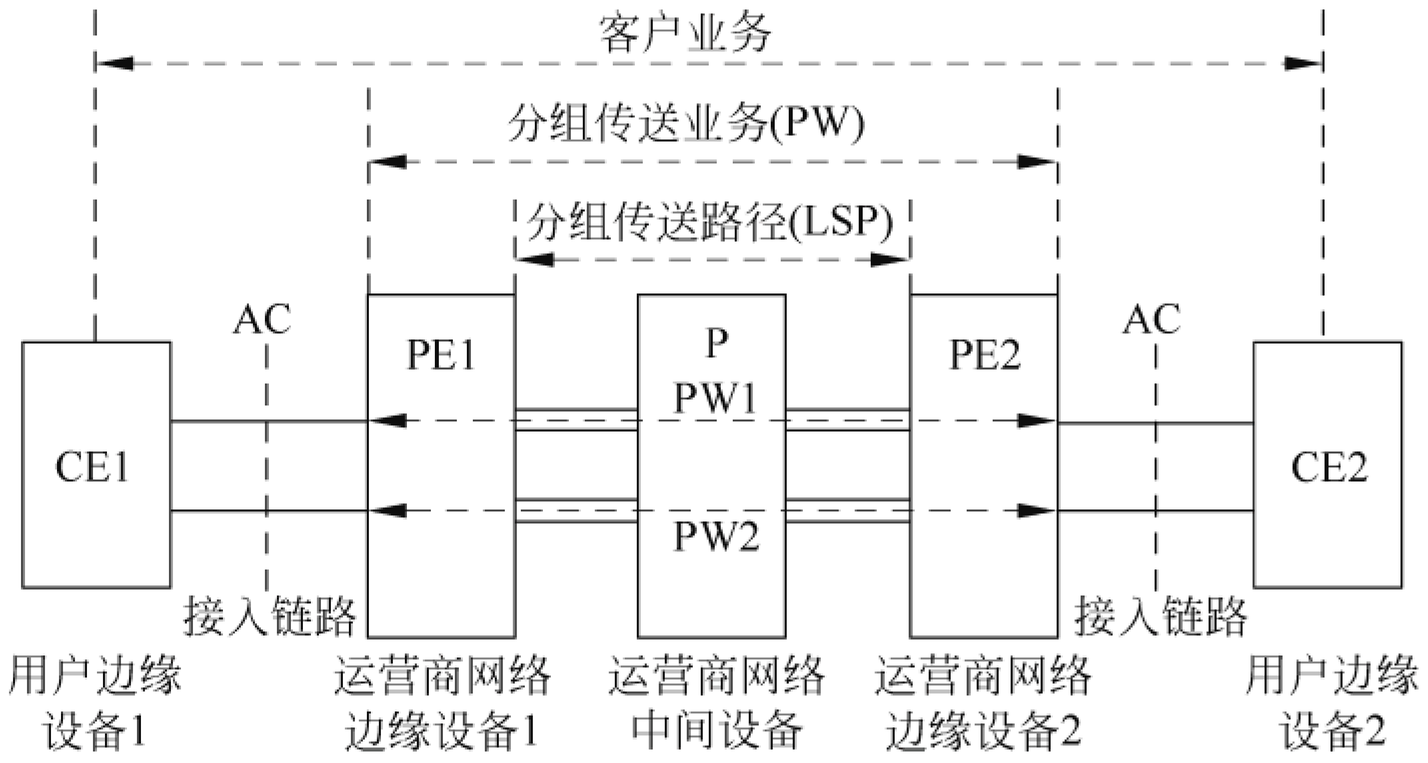


图 1.15　PTN 网元的逻辑分类

3．PTN 业务传输模型

NNI 一般为网络中两个 PTN 设备之间的接口。图 1.16 给出了 PTN 的业务传输模型，以下对各部分进行说明。

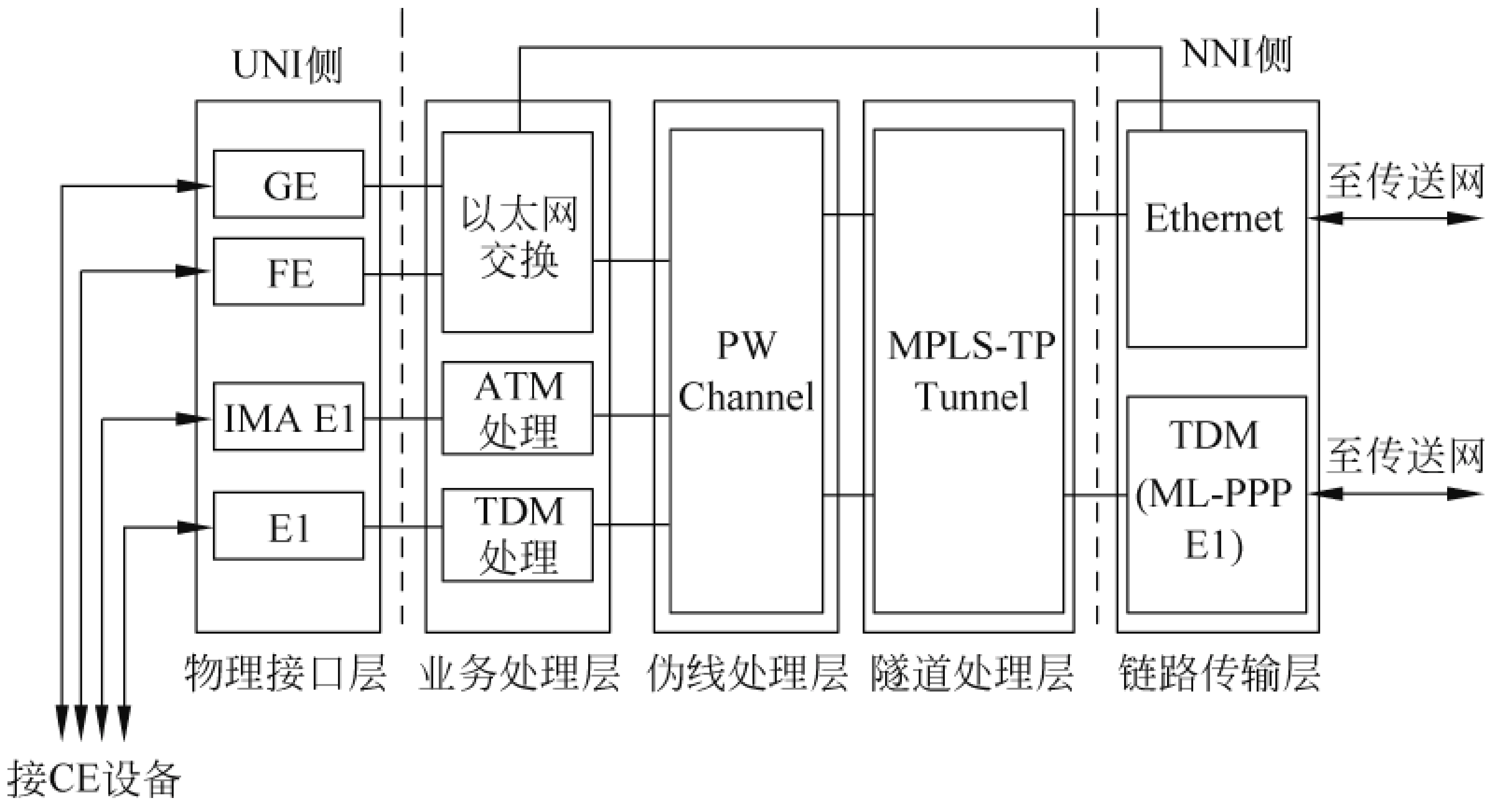


图 1.16　业务传输模型

(1) 物理接口层。当物理层采用 TDM 时，通过多链路点对点协议(Multilink-PPP，ML-PPP)接入，它是作为 PPP 功能的扩展协议。通过 ML-PPP，路由器和其他访问设备可以合并多条 PPP 链路到一个逻辑数据管道。PPP 在网络部署方面存在局限性，即一次只处理一条链接，而 ML-PPP 则不受该限制，可以将多个 PPP 链路进行捆绑，增加设备间传输的可用带宽。

当物理层为以太网络时，物理接口接收到以太网数据信号，提取以太网帧，区分以太网业务类型，并将帧信号发送到业务处理层的以太网交换模块进行处理。

接收方向：物理接口层接收由用户设备送来的物理信号(电信号或光信号)，提取信号信息，并区分业务类型后，发往相应的业务处理层进行处理。

发送方向：物理接口层接收由业务处理层送来的业务信号，根据信号类型，选择物理通道类型，并转换成在传输介质上传输的信号后，通过物理接口发往用户设备。

(2) 业务处理层。业务处理层根据不同的业务类型和业务规则，对不同业务进行相应处理。它采用边缘到边缘的伪线仿真(Pseudo-Wire Emulation Edge to Edge，PWE3)技术，支持以 TDM、ATM/IMA、FE、GE 等多种形式接入业务。

(3) 伪线处理层。对客户报文进行伪线封装(包括控制字)，并提供承载各种仿真后业务数据的方法。针对不同的仿真业务统一封装成报文格式为 PWE3 的仿真客户信号特征，并指示连接特征。从 PWE3 报文中解封装，恢复出不同的仿真业务。

(4) 隧道处理层。对 PW 进行隧道封装，完成 PW 到隧道的映射，提供分组业务转发的路径。一条隧道可承载多条伪线，通过 PW 标签区分 MPLS-TP 隧道内的不同伪线。

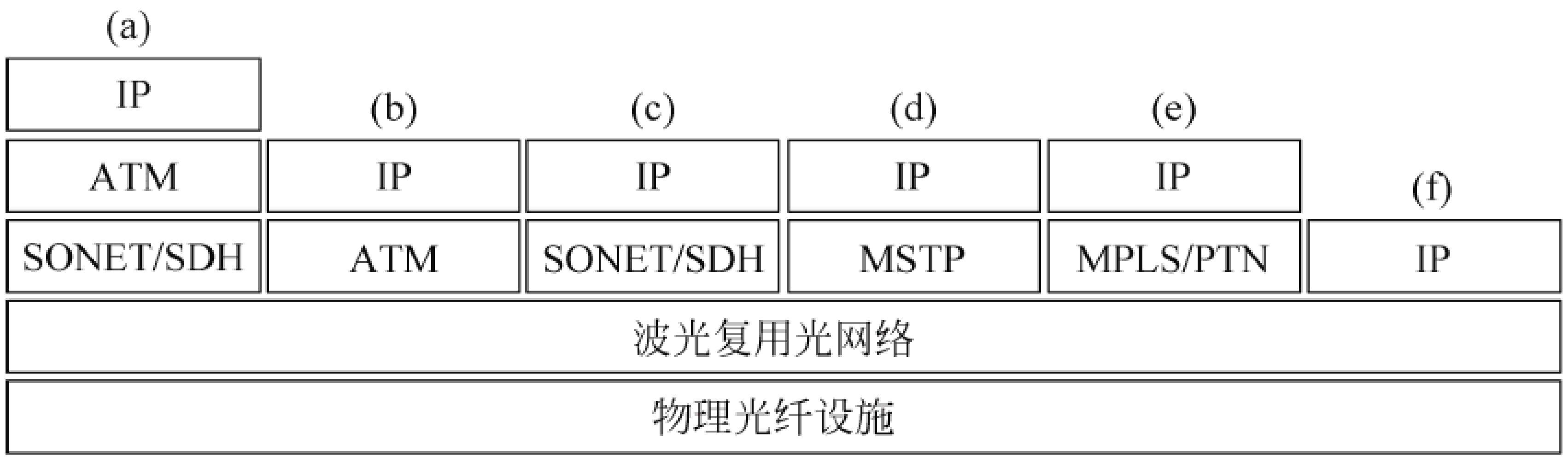
(5) 链路传输层。提供分组业务转发的路径，主要针对以太网或 TDM(E1)业务进行分组转发。

4. PTN 业务传输能力

PTN 系列分组传输设备是通过 E1、STM-1、FE、GE、10GE、40GE、100GE 等丰富的业务接口，实现 2G/3G/LTE/大客户专线等各类业务的统一接入。采用 PWE3 仿真技术，实现对各种业务的统一承载，目前的 PTN 系列分组传输设备，可支持单接口 100GE 和 40GE 等，同时还具备大规模组网的能力，完全可以满足多业务时代组建大型网的需求。

1.5.4 承载 IP 传输网演进技术

如图 1.17 所示给出了各种承载 IP 传输网络的技术演进。



图中：(a) IP over ATM (b) IP over ATM (c) IP over SONET/SDH或POS(Packet over SDH)
(d) IP over MSTP (e) IP over MPLS/PTN (f) IP over WDM

图 1.17 IP over what 的技术演进

ATM 是装在 SDH 帧中,以时分复用的方式占有部分 SDH 信道,为了在 ATM 网上能传递 IP 数据包,就发展了在 ATM 上的多协议(Multi-Protocol Over ATM,MPOA)、IP 交换,如图 1.17(a)所示,这种方法不足之处是开销大。当 ATM 能直接使用光纤时,可以在光纤上直接运行 ATM 网,如图 1.17(b)所示。由于高速路由器、IP 交换机的出现,就可以省去 ATM,直接在 SDH 等光网络上运行 IP 数据包,IP over SDH,如图 1.17(c)所示。在 SDH 的基础上,增加路由等数据接口,升级为 MSTP(多业务传输平台),即 IP over MSTP,如图 1.17(d)所示。还有 SDH 结合 MPLS(多协议标记交换)技术构成符合规范的、新型的 PTN(分组传输网),即 IP over PTN,如图 1.17(e)所示。随着波分复用(WDM)技术的日益成熟,将 SDH 时分复用的信道改为波分复用,去掉 SDH,采用 G 位(Gigabit per second)、T 位(Terabits per second)的路由交换机进行路由交换,实现真正意义上的 IP 光网络(或光因特网),即 IP over WDM,如图 1.17(f)所示。实际上,WDM 在传输网或 SDH 组网中已得到了较好的应用。

1.6 网络接口与介质

在考虑路由器或交换设备的连接接口时,主要体现在机械特性、电气特性、常用控制信号、传输速率、传输距离和接口电缆共 6 个方面。ITU-T 制定的 V 系列建议是为电话网上的数据通信而制定的,包括如下几部分:V.1～V.7(总则);V.10～V.33(接口与话音频带调制解调器);V.35～V.37(宽带调制解调器);V.40～V.42(差错控制);V.50～V.57(传输质量与维护);V.100、V.110、V.120、V.230(与其他网互通)。机械规程包括对接口的物理管脚数目、排列定义以及标准尺寸等方面的定义。

1.6.1 广域网接口与介质

1. 广域网的类型与接口

广域网是一种跨越地域的网络,目前有多种公共广域网络,用户各自的计算机网络会通过服务商提供的具体业务连接到广域网络上。

1) 广域网的类型

广域网按其提供业务的带宽不同,可简单分为窄带广域网和宽带广域网两大类。

(1) 现有的窄带公共网络包括 PSTN(公共交换电话网)、ISDN(综合数字业务网)、DDN(数字数据网)、X.25 网、Frame Relay(帧中继)网等。其中:

PSTN 主要提供电话和传真业务,通过调制解调器可以完成一些有限的数据传输业务;

ISDN 主要提供 2B+D,其中,B 表示 64Kbps、D 表示 16Kbps;

DDN 是一种广泛使用的基于点对点连接的窄带公共数据网络;

X.25 网是一种国际通用的标准广域网,基于分组交换技术,X.25 目前速率为 64Kbps～2Mbps。X.25 网络在传输数据时,沿途每个节点都要重组包,使得数据的吞吐率很低。

Frame Relay 是一种应用很广的服务,采用 E1 电路,速率可从 64Kbps～2Mbps,中间节点的延迟比 X.25 网小得多。帧中继的帧长度可变,可以方便地适应 LAN 中的任何包或帧。

(2) 现在有的宽带网络主要有 ATM 和 SDH,以及在此基础上的集成。

ATM 是异步传输模式,为在交换式 WAN 或 LAN 骨干网以及高速传输数据提供了通

用的通信机制,它同时支持多种数据类型(话音、视频、文本等)。

SDH 是目前应用最广的光传输网络,带宽高,抗干扰性强,可扩展性较强。

在具体的网络连接上一般服务商将线缆布放到用户处,用户将该线缆与一个调制设备,如普通调制解调器、数据传输单元或基带调制解调器(Data Transfer Unit,DTU)等相连,然后用户网络边缘设备(一般为路由器)再与调制设备相连。但有些线路路由器内部设有内置的设备,可以直接连接服务商布放的线缆,如 E1/ATM 等。

多业务传送平台(MSTP)可以将传统的 SDH 复用器、数字交叉链接器(DXC)、WDM 终端、网络交换机和 IP 边缘路由器等多个独立的设备集成为一个网络设备,即基于 SDH 技术的 MSTP,进行统一控制和管理。基于 SDH 的 MSTP 最适合作为网络边缘的融合节点支持混合型业务,特别是以 TDM 业务为主的混合业务。

MSTP 的特点:可以提供 10/100/1000Mbps 系列接口,通过 VC 的捆绑可以满足用户的需求;可以根据业务的需要,工作在接口组方式和 VLAN 方式;可以工作在全双工、半双工和自适应模式下,具备 MAC 地址自学习功能;可以进行 QoS 设置,QoS 的配置就是规定各接口在共享同一带宽时的优先级及所占用带宽的额度;可以对每个客户独立运行生成树协议。

2) 异步串口与同步串口

广域网按照线路类型分为 X.25、帧中继、ATM、ISDN 等,路由器因此也相应地有同/异步串口、ATM 接口、ISDN BRI 接口、PCM E1/PRI 接口等。通常中低端路由器的 WAN 接口包括异步串口、同步串口、ISDN BRI 接口及 PCM E1/PRI 接口。

(1) 异步串口。两种异步串口,一种接口名称为 Serial,将同/异步串口设置为工作在异步方式;另外一种接口名称为 Async,是专用异步串口。异步串口可以设为专线方式和拨号方式。在应用中更常用的是拨号方式,异步串口外接 Modem 或 ISDN 终端适配器(Terminal Adapter,TA)时可以作为拨号接口使用,封装链路层协议 SLIP 或 PPP,支持 IP 和 IPX 等网络协议。

(2) 同步串口。同步串口可以工作在 DTE 和 DCE 两种方式,一般情况下,路由器的同步串口工作在 DTE 方式,接收 DCE 设备提供的时钟。同步串口可以外接多种类型电缆,如 V.24 和 V.35 等,支持的链路层协议包括 PPP、帧中继、LAPB 和 X.25 等,支持 IP 和 IPX 网络层协议。

2. V.35 接口

以下简单介绍路由器的常用接口 V.35 有关规程。

(1) V.35 的特性。V.35 是一个 34 针插头/座,接口线采用平衡绞合多线对电缆。

(2) 传输速率与距离。表 1.1 给出了 IEEE 提供的 V.35 电缆在同步工作方式下以各种波特率传输数据的标准传输距离,在实际情况中,由于使用环境的差别,其传输距离的极限不尽相同。虽然 V.35 的常用速率范围为 48～64Kbps,但是它也可以支持更高的速率,如 ISDN、ATM 以及帧中继。V.35 在 100Kbps 的情况下,电缆的理论长度可以达到 1200m,实际长度要根据设备和电缆质量来确定,电缆传输(同步方式下)的最高速率是 2048Kbps。

表 1.1　传输速率与距离

V. 24		V. 35	
波特率/bps	最大传输距离/m	波特率/bps	最大传输距离/m
2400	60	2400	1250
9600	30	9600	312
64 000	20	64 000	50
115 200	10	2 048 000	30

(3) 接口电缆。V. 35 电缆一般只用于同步方式传输数据,可以在接口封装 X. 25、帧中继、PPP、SLIP、LAPB 等链路层协议,支持网络层协议 IP 和 IPX。V. 35 电缆通常用于路由器与基带 Modem 的连接之中,路由器总是处在 DTE 侧。DTE 端为 34 针型插头,DCE 端为 34 孔型插头。目前,大多数的服务单元,如分组交换机、路由器、远程网桥和网关都带有 V. 35 接口。

3. V.24 接口

V. 24 接口可工作在同步和异步两种方式下,异步工作方式下最高传输速率是 115 200bps;同步工作方式下最高传输速率是 64 000bps。

有些路由器使用 V. 24 接口,电缆可以工作在同步和异步两种方式下,所以既可以与普通的模拟 Modem、ISDN 终端适配器等以拨号方式进行异步连接,也可以连接基带 Modem 进行同步连接。异步工作方式下,封装链路层协议 PPP 支持网络层协议 IP 和 IPX,最高传输速率是 115 200bps;同步方式下,可以封装 X. 25、帧中继、PPP、HDLC、SLIP 和 LAPB 等链路层协议,支持 IP 和 IPX,而最高传输速率仅为 64 000bps。

V. 24 电缆在同步工作方式下的最大传输速率为 64 000bps;异步工作方式下,最大传输速率为 115 200bps。表 1. 1 给出了 IEEE 提供的 V. 24 电缆异步方式下以各种波特率传输数据的标准传输距离。符合 V. 24 规程的接口及电缆在通信、计算机系统中使用的非常广泛,在路由器上,主要出现在 WAN 广域网接口、AUX 备份接口和 Console 控制台接口等接口电缆之中。

4. ISDN BRI 接口

数据通信网中的有些路由器可以提供 BRI 接口,相当于内置了一个终端适配器(Integrated Services Digital Network,ISDN),可以方便地通过 ISDN 专线与远端进行通信,也可以作为一个桥梁将一个局域网接入到 Internet。路由器中包括 S/T 接口、U 接口,分别适应不同电信网络的规范。连接时,S/T 口的路由器需通过一个 NT1 再与 ISDN 线相连,U 口则可以直接连接到 ISDN 线路中使用。BRI 接口(S/T)也使用 RJ-45 水晶头连接器,U 口不通过 NT1,直接与两芯的 ISDN 用户线相连。

BRI 接口规程定制的带宽为 2B+D,B 通道的速率为 64Kbps,D 通道的速率为 16Kbps。

ISDN BRI 接口缺省封装链路层协议为 PPP,支持 IP 和 IPX 等网络层协议。

1.6.2　以太网接口与介质

以太网是一种基于总线型拓扑结构的网络,使用分布式仲裁机制来解决冲突。速度主要有 10Mbps、100Mbps 和 1000Mbps,线缆主要有双绞线、同轴电缆等。如 10Mbps 以太网

使用的线缆有 10Base-T 双绞线、10Base5 粗同轴电缆以及 10Base2 细同轴电缆。

1. 双绞线

双绞线(twisted pair)由两条相互绝缘的铜线(一般是 22～26 号绝缘铜线；线径为 0.4～1.4mm)在一起构成，这样可以使各线之间电磁干扰最小，每对线使用不同的颜色，以利于区分，并在外面包上塑料或胶皮构成对称电缆。

双绞线用来传输模拟信号和数字信号。传输模拟信号，最多的是用在电话系统中，传输距离可达 1～5km；传输数字数据信号，多用在局域网中，若速率为 100Kbps，传输距离可达 1km。若距离较近，速率可以达到 10Mbps。若采用特殊技术，速率可达 100Mbps。

双绞线分为以下两类。

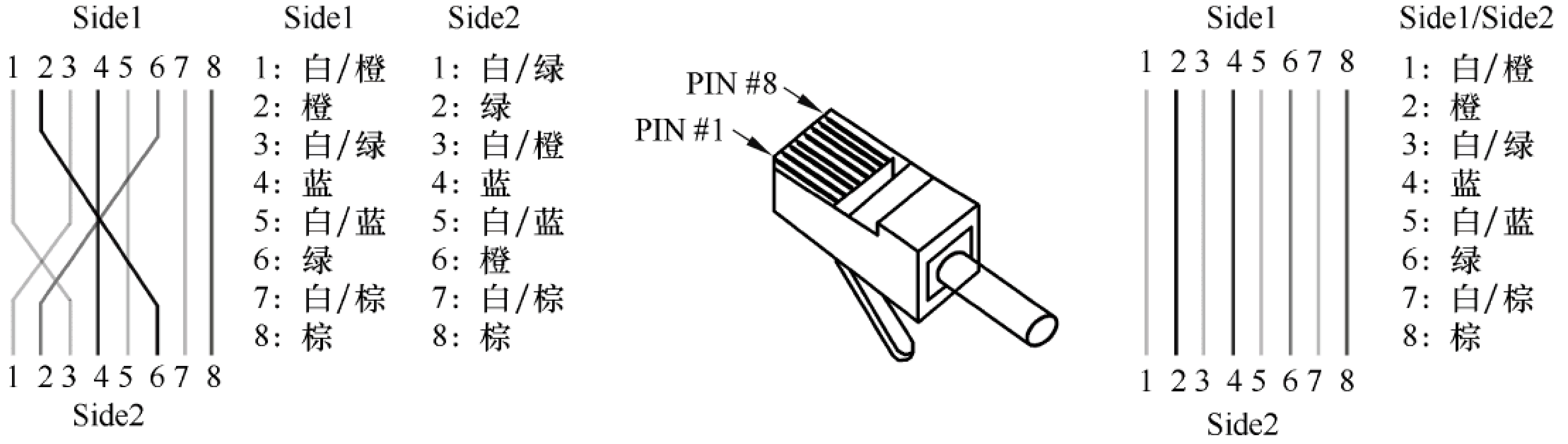
(1) 无屏蔽双绞线(Unshielded Twisted Pair，UTP)。这种双绞线没有起屏蔽作用的网状金属线。以前双绞线电缆有 5 种质量级别：1、2 类线是语音和低速数据线，带宽不大于 4Mbps；3 类线是数据线，带宽为 10～16Mbps；4 类线是数据线，带宽不大于 20Mbps；5 类线是高速数据线，带宽不大于 100Mbps；用得最多的是超 5 类线或 6 类线，带宽不小于 100Mbps。

在计算机的局域网中常用 3 类线(封皮印有 CAT 3 标记)，当传输速率为 10Mbps 时，传输距离可达 150m；5 类线(封皮印有 CAT 5 标记)，可支持 155Mbps 的数据传输。无屏蔽双绞线误码率为 10^{-6}～10^{-5}，由 2 对或 4 对双绞线组成。双绞线两端使用 RJ-45 插头，上面有 8 个槽，两端插头分别插在计算机和墙上的插线盒中。

(2) 屏蔽双绞线(Shielded Twisted Pair，STP)。这是在双绞线外面再包上一层网状金属线，用做屏蔽。使用屏蔽双绞线必须配有支持屏蔽功能的特殊连接器和相应的安装技术。传输距离在 100m 以内，速率可达 500Mbps。通常使用的速率很少超过 155Mbps，常用速率为每秒几十兆位，最大使用距离限制在几百米之内，误码率为 10^{-8}～10^{-6}。

3 类到 6 类双绞线在塑料外壳内均有这样的 4 对线缆，区别主要在于类数越高的双绞线，单位长度内的绞环数越多，拧得越紧，这使得 5 类或 6 类双绞线的交感更少并且在更长的距离上信号质量更好，更适用于高速计算机通信。

双绞线连接器一般都使用 RJ-45 连接器，如图 1.18(b)所示。5 类双绞线由 8 芯细线组成，利用细线外绝缘层上的颜色进行分组标识。通常利用单色和单色加上白色作为成对标识，也有利用色点成对进行标识的。交叉网线(又叫级连网线)和直连网线的线序如图 1.18(a)、(c)所示。图中各线画法仅为说明两端的线序关系，实际双绞电缆线芯都是成对绞在一起的。



(a) 交叉网线　　(b) 水晶头　　(c) 直连网线

图 1.18　RJ-45 连接器与双绞线的连接

目前，应用于各种网络设备的接口可能使用双绞线接口或光纤接口。双绞线和光纤接口之间不能直接相连，必须使用光电转换设备。同种线缆连接情况下各种网络设备的连接方式可以参考表 1.2。表中，N/A：又作 NA，Not Available，不可连接；normal：正常连接；cross：交叉连接。

表 1.2 设备连接方式

	主机	路由器	交换机普通接口	交换机级连口	交换机光口
主机	cross	cross	normal	N/A	SC/ST
路由器	cross	cross	normal	N/A	SC/ST
交换机普通接口	normal	normal	cross	normal	N/A
交换机级连口	N/A	N/A	normal	N/A	N/A
交换机光口	SC/ST	SC/ST	N/A	N/A	SC/ST

2. 同轴电缆及光纤

同轴电缆由两根导体组成，一根导线在中心，中间为绝缘电介材料，外面包一层同心圆筒形导体。同轴电缆是一种高带宽、低误码率、性能价格比高的传输介质，广泛用于 LAN、CATV 中。同轴电缆分为粗缆（直径为 10mm）和细缆（直径为 5mm）。除按直径划分外，同轴电缆有多种规格。

（1）基带同轴电缆（baseband coaxial cable），阻抗为 50Ω，只用于基带数字信号的传输。基带同轴电缆又分为：细缆（常用 RG-58），传输距离较近，连接时需要使用 T 形接头连接电缆和计算机中的网卡；粗缆（常用 RG-11），抗干扰性能好，传输距离远，粗缆需要使用外部收发器将信号传入计算机或发向网络。

由于计算机产生的数字数据信号不适合直接在信道中进行传输，所以在传输前仍然要经码型变换，变为曼彻斯特码后再进行传输。当传输距离不超过 1000m 时，传输速率可达 10Mbps，误码率为 10^{-11}～10^{-7}，最大传输距离限制在几千米内。

（2）宽带同轴电缆（CATV），这是电视电缆，阻抗为 75Ω，CATV 电缆可用于模拟信号或数字信号传输。当用于模拟信号传输时，频率可达 300～400MHz。要把数字信号变成模拟信号在 CATV 电缆上传输，必须采用频带 Modem。这样对于带宽为 400MHz 的 CATV 电缆，典型的数据速率为 100～150Mbps。利用频分复用（FDM）技术，把整个带宽划分为多个独立信道，分别传送数字信号、语音信号和视频图像，可以实现多种业务的综合传输，传输距离可达几十千米。当直接传输数字信号时，速率可达 50Mbps。

（3）以太网同轴电缆及光纤

10Base5 粗同轴电缆采用插入式分接头，工作速率为 10Mbps，采用基带信号。对于使用粗缆的以太网，每个干线段的长度不超过 500m，可以用中继器连接两个干线段，以扩充主干电缆的长度。每个以太网中最多可以使用 4 个中继器，连接 5 段干线段电缆。

10Base2 细同轴电缆接头采用工业标准的 BNC 连接器组成 T 形插座，使用灵活可靠性高，价格也较便宜。对于使用细缆的以太网，每个干线段的长度不能超过 200m，可以用中继器连接两个干线段，以扩充主干电缆的长度。每个以太网中最多可以使用 4 个中继器，连接 5 个干线段电缆。

表 1.3 给出了各类以太网传输距离表，在网络设计中可以提供参考。

表 1.3　各类以太网传输距离表

线缆标准	线缆类型	以太网接口	网段长度
10Base-T	双绞线	10Mbps	100m
10Base5	粗同轴电缆	10Mbps	
10Base2	细同轴电缆	10Mbps	
100 Base-TX	EIA/TIA5 类以上(UTP)双绞线 2 对	100Mbps	100m
100 Base-T4	EIA/TIA3 类、4 类、5 类(UTP)双绞线 4 对	100Mbps	100m
100 Base-FX	单模光纤(SMF)线缆	100Mbps	2～15km
100 Base-FX	多模光纤(MMF)线缆	100Mbps	550m～2km
1000 Base-T	铜质 EIA/TIA5 类以上(UTP)双绞线 4 对	1000Mbps	100m
1000Base-CX	铜质屏蔽双绞线	1000Mbps	25m
1000 Base-LX	单模，9μm，使用波长为 1300nm 激光	1000Mbps	2～15km
1000 Base-SX	多模，50/62.5μm，使用波长为 850nm 激光	1000Mbps	550m/275m
1000 Base-F	多模光纤	1000Mbps	500m

快速以太网的速度是通过提高时钟频率和使用不同的编码方式获得的。其传输方案最常用的是 100Base-T、100Base-T、100Base-TX 和 100Base-T4。100Base-T4 是一种 3 类双绞线方案，不支持全双工；100Base-TX 用得较多，使用 5 类以上双绞线，时钟信号处理速率高达 125MHz；100Base-FX 使用一对多模或者单模光纤，使用多模光纤的时候，计算机到集线器之间的距离最大可达 2km，使用单模光纤时最大可达 10km。

吉比特以太网使用 1000Base-X(8B/10B)编码，可支持 3 种介质：光纤、使用 4 对线的 5 类 UTP(1000BASE-T)和特殊的两对线 STP 电缆，也称短铜跳线(short copper jumper)。

1000Base-X 支持 3 种光纤：50μm 多模光纤、62.5μm 多模光纤和 9/10μm 单模光纤。

1000Base-X 支持两种用于激光驱动器的光波长：短波 850nm，称为 1000Base-SX；长波 1300nm，称为 1000Base-LX。

习题

一、单项选择题

1. 对于一个物理网络，数据的最大传输单元是由(　　)决定的。

   A. 硬件　　B. 软件　　C. 网管　　D. 协议

2. 在当前的数据通信网络中，存在下列哪种交换方式？(　　)

   A. 单工通信方式，半双工通信方式，全双工通信方式

   B. 基带传输方式，频带传输方式，数字数据传输方式

   C. 电路交换方式，分组交换方式，帧中继方式、信元交换方式

3. 分组交换方式与电路交换方式相比，分组交换方式的优点是(　　)。

   A. 加快了传输速度　　B. 控制简单、可靠性增加

   C. 实时性好　　D. 提高了线路的有效利用率

4. 每秒传输二进制码元的个数称为(　　)。
A. 码元速率　　B. 传送速率　　C. 数据传信率　　D. 波特率
5. 采用(　　)交换方式时,在通信进行的过程中,通信信道由参与通信的用户独享。
A. 虚电路　　B. 数据报　　C. 电路　　D. 信元
6. 协议是(　　)为了完成本层的功能而必须遵循的通信规则和约定。
A. 不同系统的对等层之间　　B. 同一系统的相邻层之间
C. 不同系统的相邻层之间　　D. 同一系统对等层之间

二、多项选择题

1. 计算机通信网可以划分为两部分,它们是(　　)。
A. 终端　　B. 通信子网　　C. 本地网　　D. 资源子网
2. 从网络覆盖范围划分,可以有(　　)这几种类型。
A. 广域网　　B. 城域网　　C. 局域网　　D. 校园网
3. 分组交换的工作方式包括(　　)。
A. 电路交换　　B. 虚电路交换　　C. 消息交换　　D. 数据报交换

三、是非判断题(将正确的题打上√)

1. 模拟信号可以转换为数字信号传输,同样数字信号也可以转换为模拟信号传输。
2. 数据通信是人—机或机—机之间的通信,必须按照双方约定的协议或规程进行通信。
3. 数据传输速率,指每秒传输的数据字节数,单位是比特/秒或 bps。
4. 局域网的传输介质通常有同轴电缆、双绞线、光纤、无线。
5. 表示层在网络需要的格式和计算机可处理的格式之间进行数据翻译。表示层执行协议转换、数据翻译、压缩与加密、字符转换以及图形命令的解释功能。
6. 会话层的主要功能是在两个节点间建立、维护和释放面向用户的连接,并对会话进行管理和控制,保证会话数据可靠传送。
7. TCP/IP 只包含了 TCP 和 IP。
8. 计算机网络的主要功能包括硬件共享、软件共享和数据传输。

四、简答题

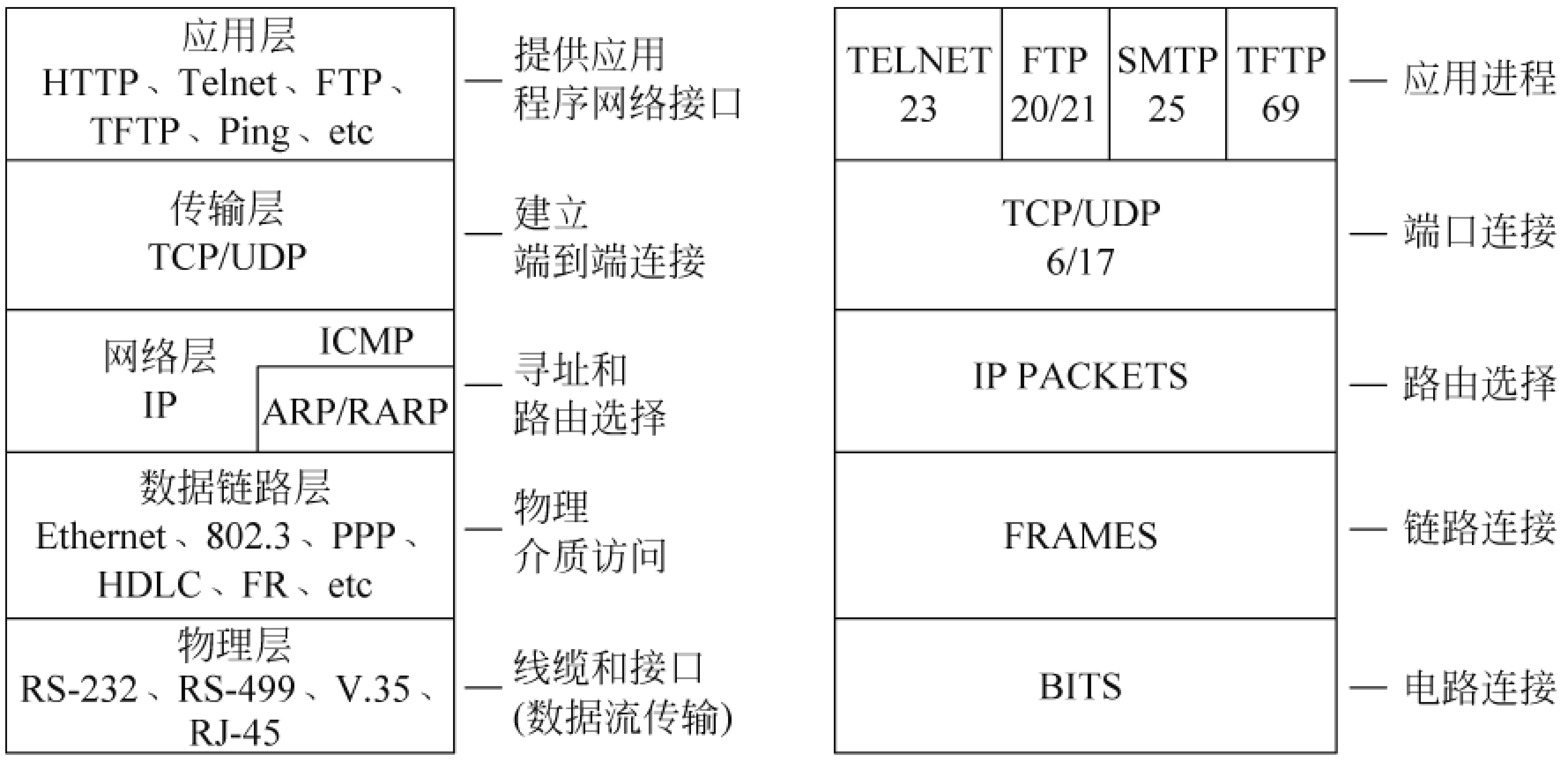
1. 物理链路和物理电路有何区别?
2. 试解释数据传输速率所使用的 3 种不同的定义(码元速率、数据传信率、数据传送速率)的主要内容。
3. 何为异步通信与同步通信?
4. 对计算机网络进行层次划分需要遵循哪些原则?
5. SDH、WDM 和 PTN 有何区别?

第2章 TCP/IP协议栈

TCP/IP 协议栈，也称互联网协议系列。TCP/IP 协议栈将 OSI 中的应用层、表示层和会话层统一划归到应用层，具体协议分布在应用层、传输层和网络层。由于 TCP/IP 可以基于不同的网络承载，因此数据链路层和物理层的位置也可以用网络接口层的位置来取代。本章将重点介绍 TCP/IP 协议栈及其对应于各层的有关协议。

2.1 TCP/IP 协议栈概述

协议栈是指在 OSI 参考模型的所有层里，相互协作或作为一个组来通信的相关协议的集合。TCP/IP 协议栈如图 2.1 所示，将 OSI 参考模型合并为 5 层：应用层、传输层、网络层、数据链路层和物理层，它同 OSI 参考模型数据封装过程一样，TCP/IP 协议在报文转发过程中，封装和去封装也发生在各对等层之间。



(a) TCP/IP协议栈　　(b) TCP/IP协议栈协议封装

图 2.1　TCP/IP 协议栈

在发送方，加封装的操作是逐层进行的。应用层将各个应用程序将要发送的数据送给传输层；传输层(TCP/UDP)将数据分段为大小一定的数据段，加上本层的报文头发送给网络层。在传输层报文头中，包含接收它所携带的数据的上层协议或应用程序的端口号，例如，Telnet 的端口号是 23。传输层协议利用端口号来调用和区别应用层各种应用程序；网络层对来自传输层的数据段进行一定的处理，利用协议号区分传输层协议，加上本层的 IP 报文头后，转换为数据包，再发送给链路层(可以是以太网、帧中继、PPP、HDLC 等)；链路层依据不同的协议加上本层的帧头，交给物理层以比特流的形式将报文发送出去。

在接收方，这种去封装的操作是从物理层一直到达应用层，逐层去掉各层的报文头部，将数据传递给应用程序执行。

TCP/IP 得到为数众多的低层协议的支持，也就是对应于 OSI 模型中的第一层(物理层)和第二层(数据链路层)。几乎流行的协议都支持 TCP/IP，例如：以太网(Ethernet)、令牌环(token ring)、光纤数据分布接口(FDDI)、端对端协议(PPP)、分组网(X. 25)、帧中继(frame relay)、异步交换模式(ATM)、同步数字传输序列(SDH)和分组传输网(PTN)等。

图 2.2 给出了计算机网络协议体系结构。在 TCP/IP 协议栈中，由于网络层以下的各层可以是不同的网络，所以也称为网络接口层。

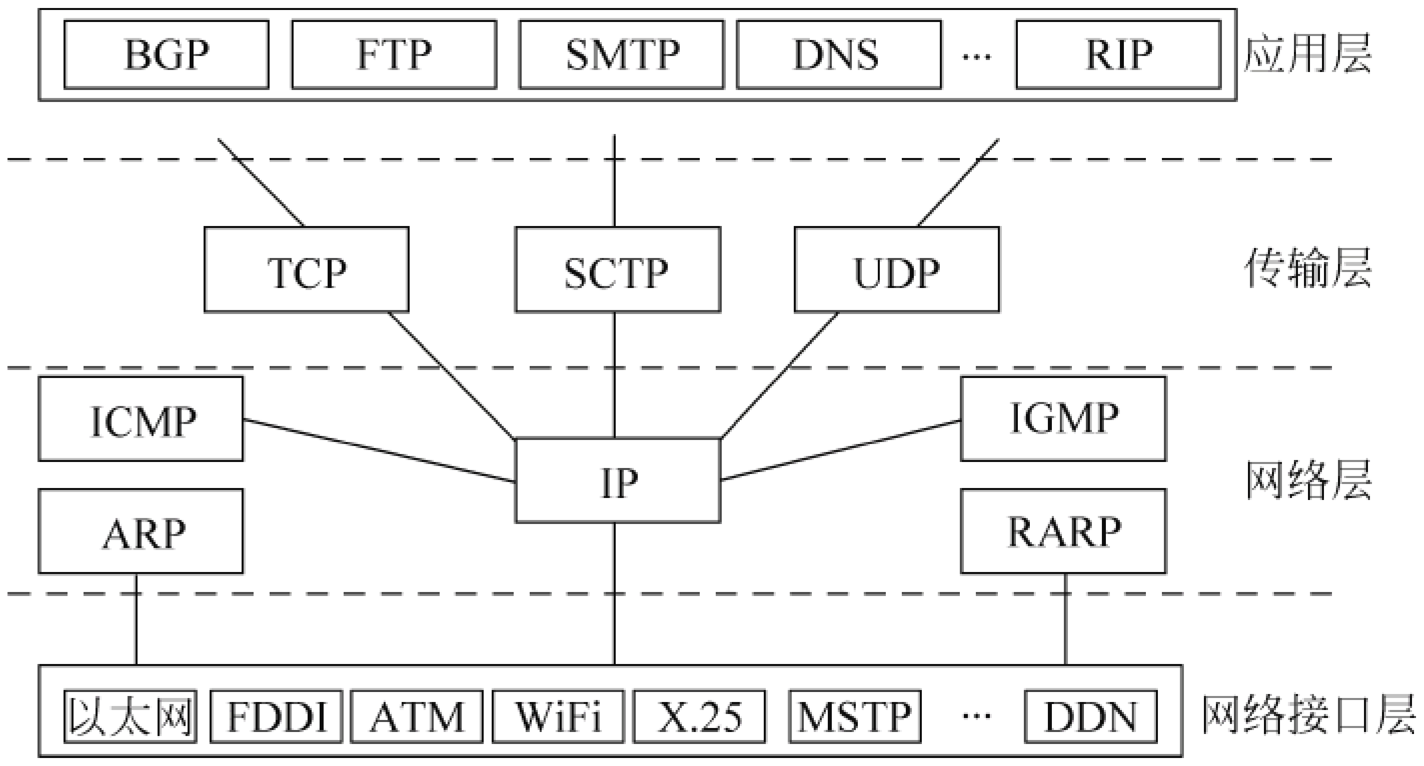


图 2.2 计算机网络协议体系结构

(1) 网络接口层：负责接收从网络层交来的 IP 数据报并将 IP 数据报通过低层物理网络发送出去，或从低层物理网络上接收物理帧，抽出 IP 数据报，交给 IP 层。网络接口有两种类型：第一种是设备驱动程序，如局域网的网络接口；第二种是含自身数据链路协议的复杂子系统，如 X. 25 中的网络接口。

(2) 网络层：也称互联层。互联协议将数据包封装成 Internet 数据包，并运行必要的路由算法。网际协议包含：负责在主机和网络之间寻址和路由的 IP 数据包协议；获得同一物理网络中的硬件主机地址的地址解析协议(ARP)；发送消息，并报告有关数据包的传送错误的网际控制消息协议(ICMP)；被 IP 主机用来向本地多路广播路由器报告主机组成员的互联组管理协议(IGMP)。

(3) 传输层：也称运送层或传送层。传输协议在计算机之间提供通信会话。传输协议的选择根据数据传输方式而定。包含：为应用程序提供可靠的通信连接的传输控制协议(TCP)，提供了无连接通信且不对传送包进行可靠的保证的用户数据报协议(UDP)；在两个端点之间提供稳定、有序的数据传递服务的流控制传输协议(SCTP)。

(4) 应用层：众多的应用程序将会通过这一层访问有关网络。

2.2 网络接口层

我们知道网络接口层可以是各种支撑 IP 数据包的网络，在这里还是要特别介绍对应 OSI 模型中的第一层(物理层)和第二层(数据链路层)，它们同属网络接口层范畴。

2.2.1 物理层

物理层是为数据通信提供透明传输的物理连接，首先要建立或激活一个连接，然后在整

个通信过程中保持这种连接,通信结束时再释放这种连接。物理层是承担数据传输的物理媒体(即通信通道),传输单位为位,它既不是只连接计算机的具体物理设备,也不是只负责信号传输的具体物理媒体,而是指在连接开放系统的物理媒体上为数据链路层提供传送比特流的一个物理连接。

1. 主要功能

物理层的主要功能就是为它服务的用户(即数据链路层)实体在具体的物理媒体上提供发送或接收比特流的能力。具体来说,其主要功能如下。

(1) 为数据端设备提供传送数据的通路。数据通路可以是一个物理媒体,也可以是多个物理媒体连接而成。一次完整的数据传输包括激活物理连接,传送数据,终止物理连接。所谓激活,就是不管有多少物理媒体参与,都要在通信的两个数据终端设备间连接起来,形成一条通路。

(2) 传输数据。物理层要形成适合数据传输需要的实体,为数据传送服务。既要保证数据能在其上正确通过,又要提供足够的带宽,以减少信道上的拥塞。

(3) 完成物理层的一些管理工作。链路层应能满足点到点、一点到多点、串行或并行、半双工或全双工、同步或异步传输等各种数据传输的需要,并能对其实施必要的管理。

(4) 通信双方从物理层建立的连接称为物理电路。一条物理电路可以支持若干条虚电路。

2. 主要特性

数据传输设备种类很多,特性各异,物理层的作用就在于要屏蔽这些差异,使得数据链路层不必去考虑物理设备和传输媒体的具体特性,而只要考虑完成本层的协议和服务。物理层上的协议有时也简称为接口特性。物理层协议规定与建立、维持及断开物理信道的有关特性,这些特性包括机械、电气、功能和规程 4 个方面。

(1) 机械特性:指明接口所用接线器的形状和尺寸、引线数目和排列、固定和锁定装置等。这很像平时常见的各种规格的电源插头,它们的尺寸都有严格的规定。

(2) 电气特性:指明在接口电缆的各条线上出现的电压范围。物理层接口的电气特性主要分为 3 类:非平衡型、新的非平衡型和新的平衡型。

(3) 功能特性:指明某条线上出现的某一电平的电压表示何种意义。

(4) 规程特性:指明对于不同功能的各种可能事件的出现顺序。

这些特性保证物理层能通过物理信道在相邻网络节点之间正确地收、发比特流信息,即保证比特流能送上物理信道,并且能在一端将它取下。物理层仅单纯关心比特流信息的传输,而不涉及比特流中各位之间的关系,对传输差错也不作任何控制。

3. 主要传输介质

目前,常用的数据信号传输介质主要有同轴电缆(coaxial cable)、双绞线(twisted pair)、光纤(fibre)、无线电波(wireless radio)等。具体在路由、交换设备的接口配置连接时包含:Console(控制线)、Coaxial(同轴电缆)、Copper Straight-through(直通线)、Copper Cross-Over(交叉线)、Fiber(光纤)、Phone(电话线)、Serial DCE、Serial DTE 等。其中 DCE 和 DTE 是用于路由器之间的连线时,需要对设置为 DCE 的路由器配置时钟。交叉线只在路由器和计算机直接相连,或交换机和交换机之间相连时才会用到。

4. **主要应用的网络**

数据通信中,物理层设备主要体现在局域网和广域网中。

(1)局域网:Xerox公司制定的以太网和IEEE 802.3标准定义了以太网物理层常用的线缆标准。其中常用的接口线缆标准有10Base-T、100Base-T、100Base-TX/FX、1000Base-T、1000Base-SX/LX。常见网络设备有中继器、集线器和交换机等,在物理层具有局域网接口。

(2)广域网:广域网物理层协议描述了DTE和DCE之间的接口。DTE指位于用户网络接口用户端设备;DCE提供到网络的物理连接口,提供了用于同步DTE和DCE设备之间数据传输的时钟信号。常用于DTE设备的有终端主机、路由器、数据输入/输出设备、通信处理机或计算机;常用于DCE设备的有广域网交换机、Modem、CSU/DSU(Channel Service Unit/Data Service Unit)。

2.2.2 数据链路层

数据链路层在物理层提供服务的基础上向网络层提供服务,TCP/IP支持多种不同的网络接口层协议,最常用的是以太网,以太网主要有两个串行接口链路层协议:串行线路网际协议(SLIP)和点对点协议(PPP)。

1. **数据链路层的概念**

为了将侧重点放在数据链路层上,如图2.3只给出了下三层的模型。

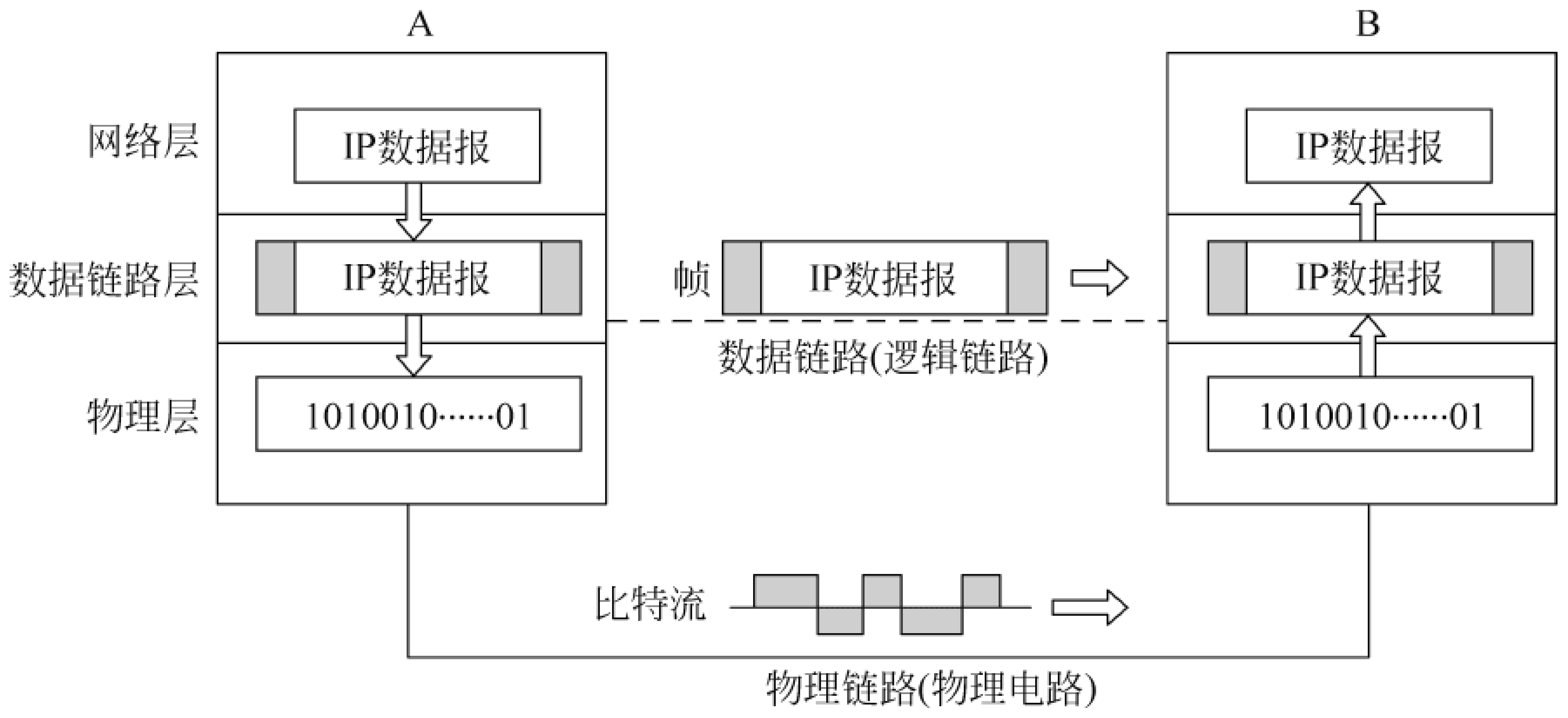


图2.3 简化的下三层模型

在图2.3中,“物理链路”和“数据链路”是两个不同的概念,物理链路(简称链路)如同物理电路,是指相邻两节点之间(期间没有任何交换节点)的一段传输线路,是从物理层看进去的。计算机网络中任意两台计算机之间的通信路径往往需要经过若干个交换节点转接,此时这两台计算机之间是由多条这样的链路串接而成的。数据链路(又称逻辑链路)则是另一个概念,它是由物理链路及实现通信协议的硬件和软件组成的。在链路上传输数据时,除了链路不可缺少,还需要有一些必要的通信协议来控制在链路上的数据传输。常用的网络适配器就是实现通信协议的硬件和软件。因此,数据链路是从数据链路层看进去的,如同一条可以在其中传输信息的数字管道,而在这条数字管道中传输的是数据帧。实际应用中的物理链路常采用多路复用技术,此时一条物理链路可以构成多条数据链路,从而提高了链路利用率。

数据链路(即逻辑链路)与链路(即物理链路)的区别在于数据链路除链路外,还必须有一些必要的规程来控制数据的传输。因此,数据链路比链路多了实现通信规程所需要的硬件和软件。由于数据链路连接具有检测、确认和重传等功能,以使不太可靠的物理链路变成可靠的数据链路,进行可靠的数据传输。

数据链路层的协议数据单元是帧。帧由首部、数据部分和尾部组成。一般来说,首部含有帧的控制信息(如地址、控制等),尾部包含帧的校验序列,数据部分作为存放网络层下传 IP 数据报的数据域。

2. 数据链路子层 LLC 与 MAC

由于广播信道可采用不同媒体的接入控制,因此在大部分网络中,数据链路层分为两个子层：逻辑链路控制子层(Logic Link Control sublayer,LLC)和介质访问控制子层(Media Access Control sublayer,MAC)。

LLC 位于网络层和 MAC 子层之间,是上层和下一层的管理层,负责流量控制、同步等,负责底层协议与网络层协议的通信。LLC 子层通过源服务访问点(Source Service Access Point,SSAP)和目的服务访问点(Destination Service Access Point,DSAP)提供了面向连接与面向无连接的网络服务环境的需要。LLC 用于管理通过单一链路连接的两个系统间的通信,它允许多个高层网络协议共享一条链路。

MAC 子层负责把物理层的“0”“1”比特流组建成帧,并且通过帧尾部的循环冗余校验(Cyclic Redundancy Check,CRC)子段进行错误检测。总之,MAC 子层定义了网络对共享介质的访问。

就像每一个人都有一个名字一样,每一台网络设备都用物理地址来标识自己,这个地址就是 MAC 地址。网络设备的 MAC 地址是全球唯一的。MAC 地址由 48 个二进制位组成,通常用十六进制数字来表示。其中前 6 位十六进制数字由 IEEE 统一分配给设备制造商,后 6 位十六进制数由各个厂商自行分配。

当有数据发送时,源网络设备查询对端设备的 MAC 地址,然后将数据发送过去。MAC 地址只适合于本网段主机的通信,另外,MAC 地址固化在硬件中,灵活性较差。

网络接口卡(Network Interface Card,NIC)有一个固定的 MAC 地址。大多数网卡厂商把 MAC 地址烧入 ROM 中。当网卡初始化时,ROM 中的 MAC 物理地址读入 RAM 中。如果把新的网卡插入计算机中,计算机的物理地址也会随之改变。

3. 主要功能与任务

1) 数据链路层的主要功能

链路管理：对于面向连接的服务,链路两端的节点在进行通信之前,发送端必须确知接收端是否处于准备接收的状态。为此,通信双方必须先交换一些必要的信息建立起这种连接,这是建立一条数据链路的过程。一旦建立起数据链路,就要维持这种连接,以确保数据传输的进行。通信完毕则释放链接。对数据链路的建立、维持和释放实施管理称为链路管理。

帧的封装和拆装：发送端的数据链路层把网络层交来的 IP 数据报加上帧首部和帧尾部封装成帧,并将其下传给物理层。接收端的数据链路层对物理层上交的帧进行差错检测,若无差错,则从收到的帧中提取 IP 数据报上交给网络层；若有差错,则将其丢弃。

帧定界：就是确定帧的边界,从传送的比特流中正确地分离出帧。数据链路层以帧为

单位传送数据。帧定界的作用就在于接收端能够从收到的比特流中准确地区分出一帧的开始和结束，即确定帧的边界位置。帧定界可采用的方法有字节填充法、位填充法、字节计数法、非法位编码法等。

透明传输：所谓透明传输是指不管链路上传输的是何种形式的位组合，都不会影响数据传输的正常进行。因为被传输的数据信息与控制信息完全一样的情况是不可避免的，此时就必须采取适当的措施，使得接收端不要将此类数据信息误认为是控制信息，从而保证数据链路层传输数据的透明性。为了解决透明传输的问题，帧界定方法采用不同的技术。

差错检测：差错检测是指在数据传输过程中用来检测是否存在差错的一种技术手段。通常，在发送的比特流后面附加差错检测码，接收端根据接收到的比特流重新计算差错检测码，然后与收到的差错检测码相比较，并根据比较结果得出是否出现差错的结论。

数据信号在通信线路上传输因受到干扰的影响，接收端往往不能接收到发送端所发送的原样比特流。误码率是衡量传输差错的度量指标，如误码率为 10^{-12}。计算机网络传输数据时也采用各种差错检测措施。循环冗余检验(CRC)是数据链路层广泛采用的一种差错检测技术。数据链路层具有“可靠传输”的功能。通常采用将差错校验的任务上交给传输层协议去处理，使数据链路层的通信效率大为提高。

2）数据链路层的主要任务

数据链路层的主要任务是提供对物理层的控制，检测并纠正可能出现的错误，使之对网络层显现一条无错线路，并进行流量调控(可选)；将网络层下传的 IP 数据报封装成帧往下传给物理层；从接收到的物理层上传无差错帧中提取 IP 数据报上交给网络层。

在 TCP/IP 协议族中，数据链路层属于网络接口层，主要负责 3 个任务：为 IP 模块发送和接收 IP 数据报；为 ARP 模块发送 ARP 请求和接收 ARP 应答；为 RARP 发送 RARP 请求和接收 RARP 应答。

4．局域网的链路层

1）IEEE 802 标准

IEEE 的数据链路层标准是当今最为流行的 LAN 标准。这些标准统称为 IEEE 802 标准。

802.1 描述了基本的局域网需要解决的问题，例如 802.1d 描述了生成树协议。

802.2 小组负责逻辑链路子层(LLC)标准的制定。

802.3 小组负责基于 CSMA/CD 访问方式的局域网络标准的制定，基于这种访问方式的网络的一个典型是 Xerox 公司发布的以太网标准。目前，我国应用最为广泛的 LAN 标准是基于 IEEE 802.3 的以太网标准。

在数据链路层常见的局域网设备有以太网交换机等。TCP/IP 协议支持多种不同的链路层协议，这取决于网络所使用的硬件，最常用的是以太网，它用到两个串行接口层协议(SLIP 和 PPP)以及环回驱动程序(loopback)。在以太网上运行 PPP 来进行用户认证接入的方式称为 PPPoE，是目前 ADSL 接入方式中应用最广泛的技术标准。最常用的以太网 IP 数据报格式是 RFC 894 封装格式。

2）SLIP

SLIP 是一种在串行线路上对 IP 数据报进行封装的简单形式，在 RFC 1055[Romkey 1988]中有详细描述。SLIP 适用于所有具有 RS.232 串行端口和高速调制解调器接入 Internet 的 PC。它是非常简单的协议，但不可避免的是其串行接口的低速率。

5. 广域网的数据链路层

广域网常见的数据链路层标准有：高级数据链路控制(High-level Data Link Control，HDLC)，是 Cisco 路由器默认的封装；点到点协议(Point-to-Point Protocol，PPP)，是华为路由器默认封装；帧中继(Frame Relay，FR)协议，是 x.25 分组交换网的改进，企业网申请帧中继时，局端(指网络运营商一端)提供数字连接识别号(DLCI)和本地管理接口(LMI)，局端是 DCE，客户端是 DTE。如局端提供的虚电路号 DLCI 是 16 和 17，本地管理类型接口 LMI 是 Cisco；综合业务数据网络(Integrated Service Data Network，ISDN)，为一系列在现有电话线上传送语音和数据的数字业务。ISDN 是一种通信协议，由电信运营商提供，由于是窄带业务，现在用得很少。

广域网常见的数据链路层设备有调制解调器(Modem)、CSU/DSU、ISDN 终端适配器、广域网交换机等。

2.3 网络层

网络层就是 OSI 参考模型的第三层，TCP/IP 协议栈的网络层利用数据链路层和物理层提供的服务来实现传输层的通信，将数据包从源网络发送到目的网络。在 Internet 中，网络层提供的协议地址或逻辑地址必须是全球唯一的。不同的网络层协议具有不同的地址格式，TCP/IP 协议栈有两种地址格式：IPv4 和 IPv6，这里主要介绍 IPv4。

2.3.1 网络层路由及其功能

网络层设备通过运行路由协议来计算到目的地的最佳路由，找到数据包应该转发的下一个网络设备，然后利用网络层协议封装数据包，利用下层提供的服务把数据发送到下一个网络设备。网络层检查网络拓扑，以决定传输报文的最佳路由，转发数据包。其关键问题是确定数据包从源端到目的端如何选择路由。与网络层路由有关的协议有被路由协议和路由协议。

被路由协议(routed protocol)，也称可路由协议，或称可被路由协议，是网络层协议，是定义数据包内各个字段的格式和用途的网络层封装协议，该网络层协议允许将数据包从一个网络设备转发到另外一个网络设备，如 IP 协议。

路由协议(routing protocol)，是应用层协议，通过在路由器之间共享路由信息来支持可路由协议。路由信息在相邻路由器之间传递，确保所有路由器知道到其他路由器的路径。路由协议创建了路由表，描述了网络拓扑结构；路由协议与可路由协议协同工作，执行路由选择和数据包转发功能。路由协议如 RIP、OSPF、BGP 等，在以后介绍。

IP 网络层的主要功能是负责相邻节点之间的数据传送。它的主要功能包括 3 个方面。

第一，处理来自传输层的分组发送请求：将分组装入 IP 数据报，填充报头，选择去往目的节点的路径，然后将数据报发往适当的网络接口。

第二，处理输入数据报：首先检查数据报的合法性，然后进行路由选择，假如该数据报已到达目的节点(本机)，则去掉报头，将 IP 报文的数据部分交给相应的传输层协议；假如该数据报尚未到达目的节点，则转发该数据报。

第三，处理 ICMP 报文：即处理网络的路由选择、流量控制和拥塞控制等问题。

网络层为了保证数据包的成功转发，主要协议有：互联网协议(Internet Protocol，IP)、地址解析协议(Address Resolution Protocol，ARP)、反向地址解析协议(Reverse Address Resolution Protocol，RARP)、网际控制消息协议(Internet Control Message Protocol，ICMP)和网际组管理协议(Internet Group Management Protocol，IGMP)。

2.3.2 IP

IP 和路由协议协同工作，寻找能够将数据包传送到目的端的最优路径。IP 提供无连接的、不可靠的服务。

1. IP 地址

IP 地址，又称逻辑地址或协议地址，它用来标识每一台网络设备或接口。在 Internet 中，为了减少路由器的路由表数目，更加有效地进行路由，清晰地区分各个网段，决定对 IP 地址采用结构化的分层方案。IP 地址给每个连接在 Internet 上的主机分配一个全球唯一的 32 位地址，用点分十进制的记法来表示 IP 地址，其地址划分类型如图 2.4 所示，从 A 类到 E 类，共为 5 类地址。IP 地址由网络地址和本网络内的主机地址两部分组成。

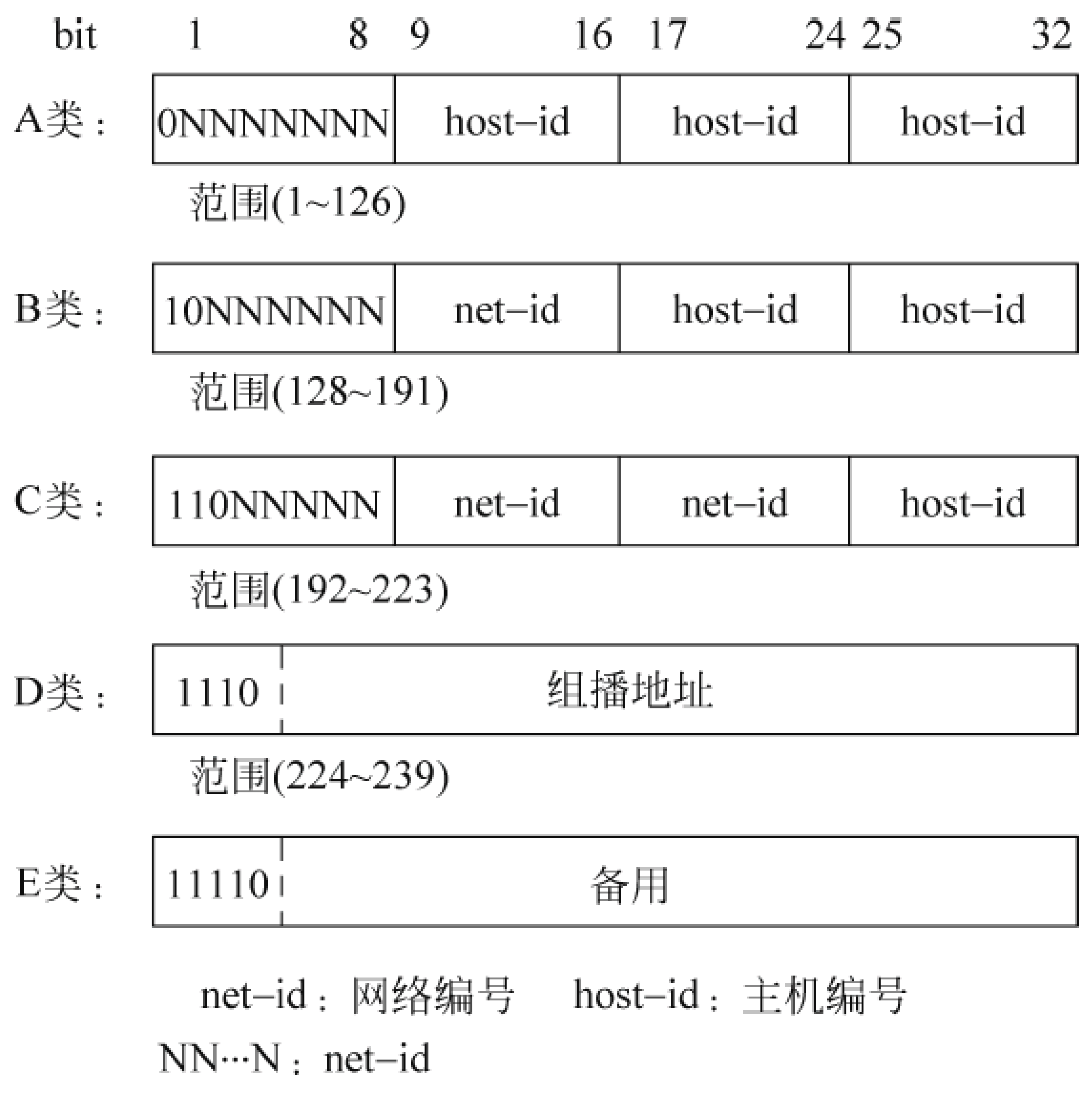


图 2.4 IP 地址类型

A 类 IP 地址的网络地址为第一个 8 位数组(octet)，第一字节以“0”开始。因此，A 类网络地址的有效位数为 8－1＝7 位，第一字节为 1～126(127 留做它用)。A 类地址的主机地址位数为后面的 3 字节，共 24 位，地址的范围为 1.0.0.0～126.255.255.255，每一个 A 类网络共有 2^{24} 个 A 类 IP 地址。

B 类 IP 地址的网络地址为前两个 8 位数组，第一字节以“10”开始。因此，B 类网络地址的有效位数为 16－2＝14 位，第一字节为 128～191。B 类地址的主机地址位数为后面的两字节，共 16 位。B 类地址的范围为 128.0.0.0～191.255.255.255，每一个 B 类网络共有 2^{16} 个 B 类 IP 地址。

C 类 IP 地址的网络地址为前 3 个 8 位数组，第一字节以“110”开始。因此，C 类网络地址的有效位数为 24－3＝21 位，第一字节为 192～223。C 类地址的主机地址部分为后面的

一字节，共 8 位。C 类地址的范围为 192.0.0.0～223.255.255.255，每一个 C 类网络共有 $2^8=256$ 个 C 类 IP 地址。

D 类地址第一个 8 位数组以“1110”开头，因此，D 类地址的第一字节为 224～239。D 类地址通常作为组播地址。常用于 X.25 和 ATM 等这类点对点的协议网络中。

E 类地址第一字节为 240～255，用于扩展和实验开发与研究。

IP 地址由国际网络信息中心(International Network Information Center，InterNIC)根据公司大小进行分配。过去通常把 A 类地址保留给政府机构，B 类地址分配给中等规模的公司，C 类地址分配给小单位。然而，随着 Internet 飞速发展，再加上 IP 地址的浪费，IP 地址已经非常紧张。

InterNIC 预留了以下网段作为私有 IP 地址：A 类地址 10.0.0.0～10.255.255.255；B 类地址 172.16.0.0～172.31.255.255；C 类地址 192.168.0.0～192.168.255.255 等。私有 IP 地址是由 InterNIC 预留的由各个企业内部网自由支配的 IP 地址。私有 IP 地址不能在公网上使用，当访问 Internet 时，需要利用网络地址转换(Network Address Translation，NAT)技术，把私有 IP 地址转换为 Internet 可识别的公有 IP 地址。但是这并不能完全解决 IP 地址短缺问题，目前较多网络已经正式启用了 IPv6 协议。IPv6 地址有 128 个二进制位，共约 2^{128} 个 IP 地址，完全可以解决 IP 地址紧张的问题。

一些特殊的 IP 地址被用于别的用途，不能用于标识网络设备，特殊 IP 地址的使用见表 2.1，表 2.2 列出了特殊 IP 地址的取值范围。常用 IP 地址的取值范围见表 2.3。

表 2.1 特殊 IP 地址的使用

net-id	host-id	源地址使用	目的地址使用	代表的意思
0	0	可以	不可	在本网络上的本主机
0	host-id	可以	不可	在本网络上的某个主机
全一	全一	不可	可以	只在本网络上广播(各路由器均不转发)
net-id	全一	不可	可以	对 net-id 上的所有主机进行广播
127	任何数	不可	可以	用做本地软件的回送测试(loopback)

表 2.2 特殊 IP 地址的取值范围

127.0.0.1	保留作为环路测试地址
主机地址全为 0	网络地址，例如，网络号为 172.16.0.0
节点地址全为 1	代表某个网段的广播地址，例如：172.16.255.255
0.0.0.0	用于 Cisco 路由器来指定默认路由
255.255.255.255	广播地址，广播网络上的所有节点
10.0.0.0 172.16.0.0～172.31.0.0 192.168.0.0～192.168.255.0	保留地址

表 2.3 常用 IP 地址的取值范围

网络类型	最大网络数	最小网络地址	最大网络地址	最大主机数(每个网络中)
A	126	1	126	16 777 214
B	16 384	128.0	191.255	65 534
C	2 097 152	192.0.0	223.256.256	254

对于 IP 地址为 127. x. x. x 和 127.0.0.1，表示回环地址和本地软件回送测试之用，例如，127.0.0.1 往往用于环路测试。

对于主机部分全为“0”的 IP 地址，称为网络地址，网络地址用来标识一个网段。

对于主机部分全为“1”的 IP 地址，称为网段广播地址，广播地址用于标识一个网络的所有主机。例如，10.255.255.255，192.168.1.255 等，路由器可以在 10.0.0.0 或者 192.168.1.0 等网段转发广播包。广播地址用于向本网段的所有主机发送数据包。

全“0”的 IP 地址 0.0.0.0，常用于代表默认网络，在路由器表中用于构造默认路径。有的路由器用 0.0.0.0 地址指定默认路由。

全“1”的 IP 地址 255.255.255.255，也是广播地址，但 255.255.255.255 代表所有主机，用于向网络的所有主机发送数据包。这样的广播不能被路由器转发。

IP 地址主要特点概括如下。

(1) IP 地址分等级(网络号＋主机号)的好处：IP 地址管理机构在分配 IP 地址时只分配网络号(最开始按类，后来分子网)。路由器仅根据目的主机所连接的网络号来转发分组，这样就使得路由表中的项目数大幅度减少，从而减小路由表存储空间。

(2) IP 地址(或其中某个字段)和主机地理位置没有对应关系。IP 地址管理机构在分配 IP 地址时只分配网络号，实质是从地址管理机构到用户逐级提供路由支持。

(3) 所有 net-id 都是平等的。同一局域网上的主机或路由器对应端口的 IP 地址中的网络号必须是一样的。

(4) 2 个路由器直接相连时，可以指明 IP 地址(最多只有 2 个)，也可以不指明。

(5) 在 IP 地址中，可以将主机字节的一部分作为子网号(subnet-id)，这样可以增加网络灵活性，减少广播流量，优化网络性能，简化管理，易于扩展网络，解决 IP 地址不足等问题。

(6) 每一个网段会有一些 IP 地址不能用做主机 IP 地址，假设本网段的主机部分位数为 n，那么可用的主机地址个数为 2^n-2 个。例如 B 类网段 172.16.0.0 有 16 个主机位，因此有 2^{16} 个 IP 地址，去掉一个网络地址 172.16.0.0，一个广播地址 172.16.255.255 不能用做标识主机，那么共有 $2^{16}-2$ 个可用地址。

2. IP 数据报的格式

IP 数据报的格式如图 2.5 所示，所有的 TCP、UDP、ICMP 及 IGMP 等协议报文都可以作为 IP 数据报格式中的数据。IP 数据报中各项的含义如下。

(1) 版本：即所使用的 IP 的版本，目前使用的是版本号为 4 的 IP，即 IPv4。

(2) 首部长度：指的是首部占 32 位的字段数，包括任选项。

(3) 服务类型(Type of Service，ToS)：字段包括一个 3 位的优先权子字段(现在已被忽略)、4 位的 ToS 子字段和 1 位未用位(但必须置 0)。4 位的 ToS 分别代表所采用的服务类型是最小时延、最大吞吐量、最高可靠性或最小费用的一种，表 2.4 给出了部分 ToS 推荐值。

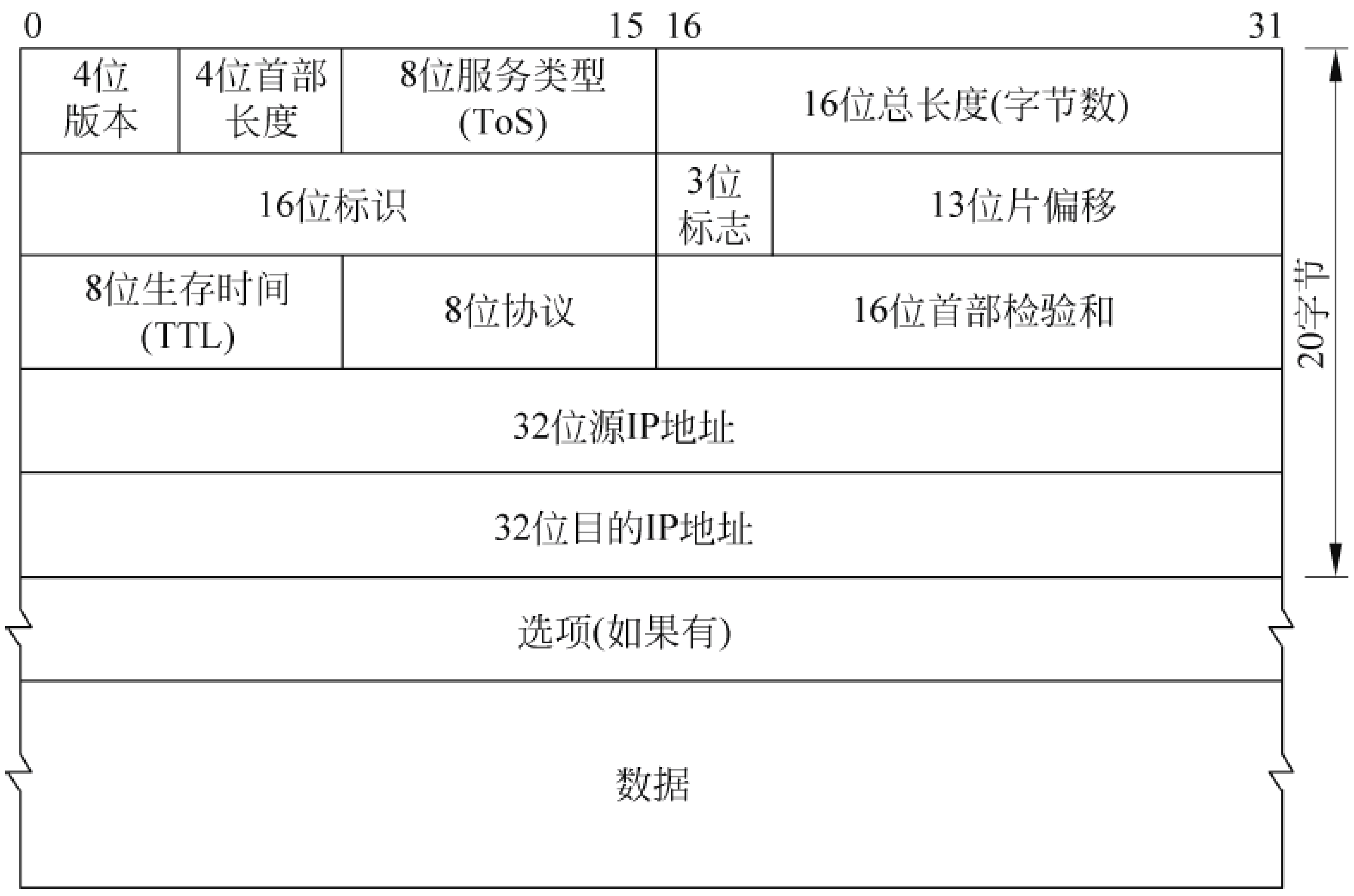


图 2.5　IP 数据报的格式

表 2.4　服务类型字段推荐值简表

应 用 程 序	最 小 时 延	最大吞吐量	最高可靠性	最小费用	十六进制
控制	1	0	0	0	0×10
数据	0	1	0	0	0×08
TFTP	1	0	0	0	0×10
SNMP	0	0	1	0	0×04

(4) 总长度：指整个 IP 数据报的长度，以字节为单位。利用首部长度字段和总长度字段，就可以知道 IP 数据报中数据内容的起始位置和长度。由于该字段长 16 位，所以 IP 数据报理论上最长可达 65 535 字节，但是大多数的链路层都会对它进行分片。而且，主机也要求不能接收超过 576 字节的数据报。事实上大多数允许超过 8192 字节的 IP 数据报。

(5) 标识：唯一地标识主机发送的每一份数据报。通常每发送一份报文它的值就会加 1。对于发送端发送的每份 IP 数据报来说，其标识字段都包含一个唯一值，该值在数据报分片时被复制到每个片中。

(6) 标志：3 位，最后一个位称为 MF(More Fragment)，用来表示“更多的片”。除了最后一片置 0 外，其他每个组成数据报的片都要把该位置 1；中间一个位为“不分片”标志 DF(Don't Fragment)，如果将这一位置 1，后续设备中的 IP 将不对数据报进行分片，如传送的数据报太长不能通过某节点时，只有丢弃并返回一个 ICMP 差错报文，只有置 0 时才允许分片。

(7) 片偏移：指的是该片偏移原始数据报开始处的位置。另外，当数据报被分片后，每个片的总长度值要改为该片的长度值，要求每一片的长度一定是 8 字节的整数倍。当 IP 数据报被分片后，每一片都成为一个分组，具有自己的 IP 首部，并在选择路由时与其他分组独立。这些分片数据报将在接收端重新组装，恢复原始数据报。

(8) 生存时间(Time To Live，TTL)：该字段设置了数据报可以经过的最多路由器数。它指定了数据报的生存时间。TTL 的初始值由源主机设置(通常为 32 或 64 等)，一旦经过一个处理它的路由器，它的值就减去 1。当该字段的值为 0 时，数据报就被丢弃，并发送

ICMP 报文通知源主机。

(9) 协议：指 IP 报文中封装的数据是来自何种协议，可以是传输层或其他层的协议，如 ICMP、IGMP、UDP 和 TCP 等协议号。

(10) 首部检验和：根据 IP 首部计算检验和码。它不对首部后面的数据进行计算。

(11) 源 IP 地址和目的 IP 地址字段：分别标明了发送端和接收端的 IP 地址。

(12) 选项：根据需要确定，并非所有主机和路由器都支持这一项。

【例 1】 某一路由器在 IP 层转发一个 2200 字节的上层报文，加上 20 字节的首部后，可形成一个标识号为 50 的 IP 数据报。但由于后面要经过局域网所能传送的最长数据帧中的数据部分只有 1120 字节。因此数据报在路由器中必须进行分片，试问局域网要向其上层总共要传送多少字节的数据?

后面局域网所能传送的最长数据帧中的数据部分只有 1120 字节，即每个 IP 数据片的数据部分要小于或等于(1120－20)字节，由于片偏移是以 8 字节(即 64 位)的整数倍计数，所以 IP 数据片的数据部分最大不超过 1024 位，这样 2200 字节的报文要分为 3 个数据片，分别是 1024 字节、1024 字节和 152 字节，分片后数据报首中有关字段的数值见表 2.5。当然局域网向其上传送的总字节数为：

$$2200 + 3 \times 20 = 2260 \text{ 字节}$$

表 2.5 分片后数据报首中有关字段的数值表

数据报片	每片总长度/字节	标识	MF	DF	片 偏 移
1	1042(1024＋20)	50	1(后面还有片)	0(允许分片)	0(0/8)
2	1042(1024＋20)	50	1(后面还有片)	0	128(1024/8)
3	172(152＋20)	50	0(最后一片)	0	256(2048/8)

2.3.3 ICMP 和 IGMP

1. ICMP

ICMP 是一个网络层的协议，定义了网络层控制和传递消息的功能。ICMP 是 Internet 控制报文协议，ICMP 被认为是 IP 层的一个组成部分。它传递差错报文以及其他需要注意的信息。ICMP 报文通常被 IP 层或更高层协议(TCP 或 UDP)使用。一些 ICMP 报文把差错报文返回给用户进程，ICMP 报文是在 IP 数据报内部被传输的。

ICMP 包含几种不同的消息，其中 ping 程序借助于 echo request 消息，主机可通过它来测试网络的可达性，ICMP Echo Reply 消息表示该节点是可达的。ICMP 还定义了源抑制(source quench)报文。当路由器的缓冲区满后，送入的报文被丢弃，此时路由器向发送报文的主机发送源抑制报文，要求降低发送速率。

差错报告是指产生错误时，ICMP 只向数据报的源发端回送差错情况，源必须把差错交给一个应用程序或采取其他措施来纠正。

ICMP 报文帧结构如图 2.6 所示，在每一层都要加一个头。ICMP 的特点是：与携带用户信息的数据具有完全相同的路由选择；不报告携带 ICMP 报文的数据报出现的误差；不具备可靠性和优越性，是不可靠的无连接服务。

ICMP 报文格式如图 2.7 所示，不同类型 ICMP 报文由类型字段(Type)和代码字段

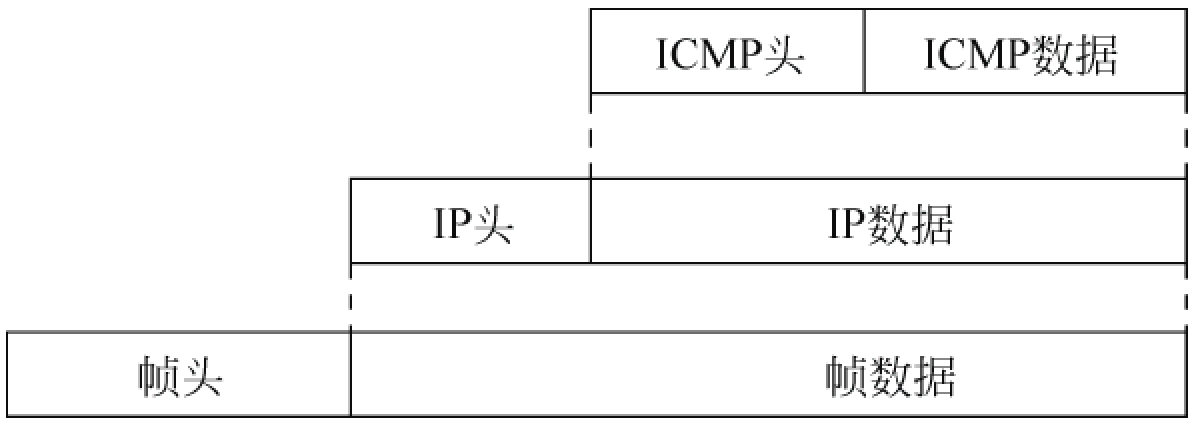


图 2.6 ICMP 报文帧结构图

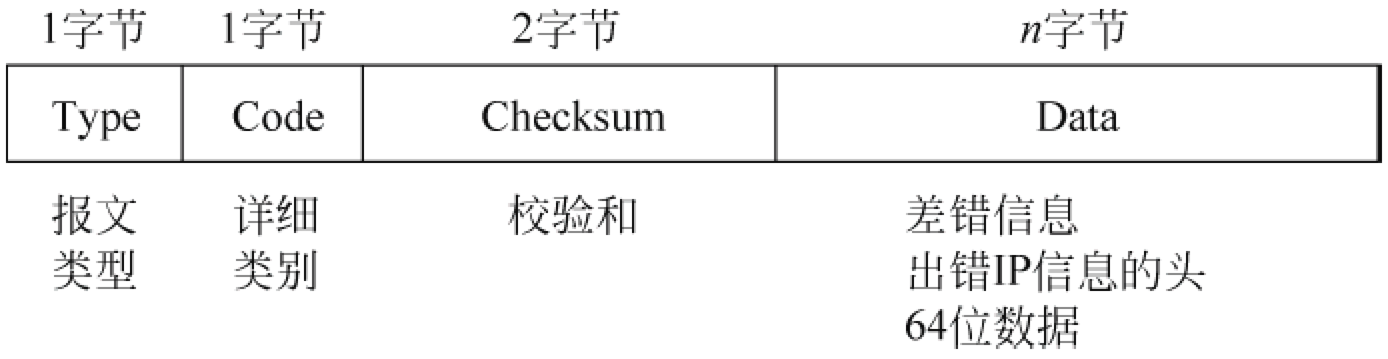


图 2.7 ICMP 报文格式

(Code)共同决定。具体定义由专表描述,表明 ICMP 报文是一份查询报文还是一份差错报文。因为对 ICMP 差错报文有时需要做特殊处理,因此我们需要对它们进行区分。例如,在对 ICMP 差错报文进行响应时,永远不会生成另一份 ICMP 差错报文。如果没有这个限制规则,可能会遇到一个差错产生另一个差错的情况,这样会无休止地循环下去。ICMP 功能概括如下。

(1) 检查目的站的可到达性与状态:主机或路由器向指定的站发送 ICMP ECHO 请求报文,请求报文包含一个可选的数据区;收到 ECHO 报文的机器应立即回应一个 ECHO 应答报文,应答报文包含了请求报文中数据的副本,如 ping。

(2) 目的端不可到达报告:向源发送一个目的端不可到达报文,并丢弃数据报。

(3) 拥塞和数据流控制:高速的计算机与低速的网络处理能力不匹配;多个计算机要同时通过一个网络的路由器等,发生拥塞的路由器为每个丢弃的数据报发送一个源抑制报文。

(4) 改变路由请求:假定路由器是知道正确路由的。

(5) 检测循环或过长的路由:一旦路由器因数据报的下一跳计数器为零或等待分段重组超时就向源发回一个 ICMP 超时报文,造成生存期到时还是超时等问题的主要原因是:出现循环路由;源和目的离得太远(超长)。

(6) 报告其他问题:当路由或主机发现一个数据报的问题时(如不正确的数据报头),便向源发端发送一个参数问题的 ICMP 报文,指针标识数据报中产生问题的字节。

(7) 时钟同步和传送时间估计值:计算请求到目的地,被转换成为应答及返回所需的时间,计算网络传送时间估计远程和本地时钟的区别。

(8) 获得子网地址的掩码:为了了解本地网络使用的子网掩码,主机可向路由器发出一个地址掩码请求(address mask request)报文,并可接收到一个地址掩码报文。

(9) 用 ICMP 报文跟踪路由(trace route 工具):利用 ICMP 超时报文发现目的地的一条路径上有路由器列表,将目的列表等信号带回到源。

2. IGMP

Internet 组管理协议,IGMP 报文能够让一个物理网络上的所有系统知道主机当前所在的多播组。多播路由器需要这些信息以便知道多播数据报应该向哪些接口转发。正如

ICMP 一样，IGMP 也被当作 IP 层的一部分。IGMP 报文通过 IP 数据报进行传输，它有固定的报文长度，没有可选数据。当 IP 首部中协议字段值等于 2 时，即表明该 IP 数据报为 IGMP 报文。IGMP 报文的具体格式如图 2.8 所示。

0　3	4　7	8　15	16　31
4位IGMP版本(1)	4位IGMP类型(1~2)	未用	校验和
32位组地址(D类IP地址)			

图 2.8　IGMP 报文的具体格式

(1) 版本字段：取 1。

(2) 类型字段：当该字段为 1 时，表示该 IGMP 消息是一个主机从属关系查询消息，它使路由器能够查询网络上多播组的成员；当该值为 2 时，表明该消息是一个主机从属关系报表消息，它使主机能够显示组中成员的关系，是对主机关系查询的响应。

(3) 组地址字段：指明多播组地址。当 IGMP 消息是主机从属关系查询消息时，该字段为 0；当 IGMP 消息是主机从属关系报表消息时，该字段中保存的是 IP 多播地址。

2.3.4　ARP 和 RARP

1. ARP(地址解析协议)

ARP 为 IP 地址到对应的硬件地址之间提供动态映射。我们之所以用动态这个词是因为这个过程是自动完成的，一般应用程序用户或系统管理员不必关心。它为两种不同的地址形式提供映射：32 位的 IP 地址和数据链路层使用的任何类型协议的物理地址。ARP 数据格式如图 2.9 所示。

硬件地址类型		协议地址类型
硬件地址长度	协议地址长度	操作(请求/应答)
发送方硬件地址(0~3字节)		
发送方硬件地址(4~5字节)		发送方IP地址(0~1字节)
发送方IP地址(2~3字节)		目的硬件地址(0~1字节)
目的硬件地址(2~5字节)		
目的IP地址(0~3字节)		

图 2.9　ARP 数据格式

将一台主机的 IP 地址翻译成等价的硬件地址的过程只限于本网络解析。一个 ARP 消息被放在一个硬件帧后，被广播给网上所有的计算机，每台计算机收到该请求后，都会检查 IP 地址，与 IP 地址不匹配的计算机会丢弃收到的请求，不发任何回答信号。

有两类 ARP 消息，分别是请求(ARP request)和应答(ARP reply)。

【例 2】 每个主机都对应一个 ARP 高速缓存(ARP Cache)，里面有 IP 地址到物理地址的映射表，如果某个主机缺少某项地址映射表，就通过自动运行 ARP 获取，假如 B 主机中没有 C 主机的地址映射表，ARP 消息传送过程如图 2.10 所示，试写出其具体步骤。

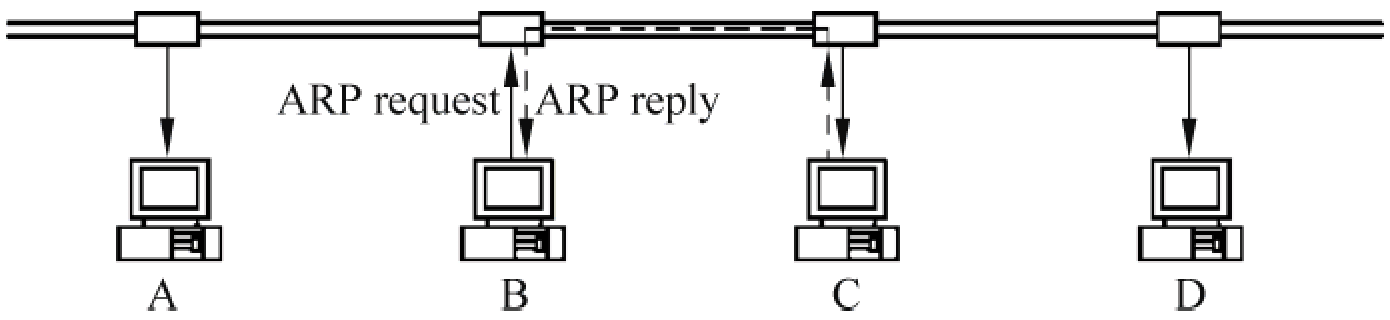


图 2.10　ARP 消息的传送过程

答：具体步骤如下：

(1) B 在本 LAN 中广播发送一个 ARP 分组，含有 C 的 IP 地址。

(2) LAN 所有主机上都收到此 ARP 分组。

(3) C 见到有自己的 IP 地址，就向 B 回送一个 ARP 相应分组，并附上自己的物理地址。

(4) B 收到后，就将 C 的 IP 地址到物理地址的映射写入其 ARP cache 中。

(5) 以后 B 就可以和 C 直接通过物理地址通信了。

2. RARP(逆地址解析协议)

RARP 是那些没有磁盘驱动器的系统使用(一般是无盘工作站或终端)的，它需要系统管理员进行手工设置。在地址转换时，RARP 使只知道自己物理地址的主机能够得到 IP 地址，并需要有一个主机充当 RARP 进程的服务器。它的实现过程是从接口卡上读取唯一的硬件地址，然后发送一份 RARP 请求(一帧在网络上广播的数据)，请求某个主机响应该无盘系统的 IP 地址(在 RARP 应答中)，在获得 IP 地址后，就能够像其他工作站一样正常地访问网络了。

对应于 ARP、RARP 请求以广播方式发送，ARP、RARP 应答一般以单播方式发送，以节省网络资源。

2.4　传输层

传输层协议规定在源节点和目的节点的两个进程实体之间，提供一种端到端的数据传输方式。本节主要介绍 TCP、UDP 和 SCTP。

2.4.1　传输层协议功能与端口

1. 传输协议功能

TCP/IP 模型最早提供传输控制协议(TCP)和用户数据报协议(UDP)，而流控制传输协议(Stream Control Transport Protocol，SCTP)是为了满足信令传输的多宿性要求而提出的一种新的传输层协议，多用在通信网络的软交换中。多宿是指一个 SCTP 可以通过多个 IP 地址到达，这样两个 SCTP 端点在建立偶联后，数据可以通过不同的物理通路进行传送。

(1) TCP：为应用程序提供可靠的面向连接的通信服务，适用于要求得到响应的应用程序。目前，许多流行的应用程序都使用 TCP。

(2) UDP：提供了无连接通信，且不对传送数据包进行可靠的保证。适用于一次传输小量数据，可靠性则由应用层来负责。

传输层的目的就是向高层提供有效、可靠的服务，主要定义了主机应用程序间端到端的

连通性，它一般包含4项基本功能。

(1) 将应用层发往网络层的数据分段或将网络层发往应用层的数据段合并。

(2) 建立端到端的连接，主要是建立逻辑连接以传送数据流。

(3) 将数据段从一台主机发往另一台主机。在传送过程中通过计算校验和流控制的方式保证数据的正确性，流控制可以避免缓冲区溢出。

(4) 部分传输层协议保证数据传送的正确性。主要是在数据传送过程中，确保同一数据既不多次传送，也不丢失。同时还要保证数据包的接收顺序与发送顺序一致。

2. 传输协议端口及套接字

TCP、UDP所提供的服务端口号(熟知端口)是在1～1023，端口号具体由Internet号码分配机构(Internet Assigned Numbers Authority，IANA)分配管理。其中，低于255的端口号保留用于公共应用；255～1023的端口号分配给各个公司，用于特殊用途；对于高于1023的端口号，称为临时端口号，IANA未做规定。TCP和UDP使用16位端口号(或者Socket)来表示和区别网络中的不同应用程序。不同协议的端口号存放在与其对应的协议报文中，一般是存放在下层协议中，也有的协议在同一层。协议号或端口号都指的是网络协议号，也称端口地址，以下给出几种常用的协议及其端口分配。

(1) TCP报文中的端口号：HTTP为80(WWW)，FTP为20/21(数据/控制)，Telnet为23，SMTP为25，DNS为53，POP3为110端口等；

(2) UDP报文中的端口号：DNS为53，BootP为67/68(server/client)，TFTP为69，SNMP为161/162等；

(3) IP数据报文中的协议字段：TCP为6，UDP为17，ICMP是1。

而IP的协议号是0，是放在下一层协议报文的类型字段中。

Socket(套接字)是网络程序间通信的一种方法。每一个套接字都用一个半相关描述：{协议，本地地址、本地端口}；一个完整的套接字则用一个相关描述：{协议，本地地址、本地端口、远程地址、远程端口}。每一个套接字都有一个由本地操作系统分配的唯一的套接字号。

Socket主要包括以下3种类型。

(1) 流式套接字(SOCK_STREAM)：流式套接字提供可靠的、面向连接的通信流。典型的有TCP，能保证数据传输的正确性和顺序。

(2) 数据报套接字(SOCK_DGRAM)：数据报套接字定义了一种无连接的服务，数据通过相互独立的报文进行传输，是无序的，并且不保证可靠、无差错。代表协议是UDP。

(3) 原始套接字(SOCK_RAM)：原始套接字允许对底层协议如IP或ICMP直接访问，它功能强大但使用不太方便，主要用于一些协议的开发。

2.4.2 TCP

1. TCP报文

TCP是一个可控的传输层协议，负责流量控制和数据报排序，在彼此交换数据之前必须先建立一个TCP连接，在一个TCP连接中，仅有两方彼此通信。所以广播和多播不能用于TCP。它的操作是由其报头中的数据来管理的，TCP把该报头添加到要传输的数据上。如图2.11所示的是TCP报文格式。

图 2.11 TCP 报文格式

(1) 16 位源端口号：指明发送数据主机上的应用程序所在的端口。

(2) 16 位目的端口号：指明接收数据主机上的对应应用程序所在的端口。

每个 TCP 报文头部都包含源端口号和目的端口号，用于标识和区分源端设备和目的端设备的应用进程。在 TCP/IP 协议栈中，源端口号和目的端口号分别与源 IP 地址和目的 IP 地址组成套接字，唯一地确定一条 TCP 连接。Socket 分为源套接字和目的套接字，可简单表述为

源套接字：源端口号＋源 IP 地址

目的套接字：目的端口号＋目的 IP 地址

(3) 32 位序号：指明从 TCP 发端向 TCP 收端发送的数据字节流。它表示在这个报文段中的第一个数据字节。如果将字节流看作在两个应用程序间的单向流动，则 TCP 用序列号对每字节进行计数。

(4) 32 位确认序号：指明接收端希望接收的下一字节的序列号。包含发送确认的一端所期望接收到的下一个序号。因此，确认序号应该是上次已成功收到的数据字节序列号加 1。

(5) 4 位首部长度：指明首部中 32 位(即 4 字节)的数目。所以 TCP 最多可以有 64 字节的首部(主要是选项字段的长度不固定)。如果没有选项字段，那么该值为 5。

(6) 6 位标志字段：从高位到低位依次是 URG(紧急指针)、ACK(确认序号有效)、PSH(接收方应尽快将报文交付应用层)、RST(重建连接)、SYN(同步序号，用来发起一个新连接)和 FIN(释放一个连接，发送端完成发送任务)。

(7) 16 位窗口大小：指明接收端准备接收的字节数目。

(8) 16 位校验和：是 TCP 报头数据的 16 位校验和。

(9) 16 位紧急指针：在标志字段中的 URG 位为 1 时，和序号字段中的值相加以识别出数据段中紧急数据的最后一字节的序号。

(10) 选项字段：用于与传输同位体通信。

(11) 数据字段：为用户数据部分。

2. TCP 通信原理

TCP 协议是可靠的、面向连接的传输。下面以客户端、服务器连接为例进行说明。

1）TCP 建立连接

在 TCP 协议中，当服务器与客户端建立连接时需要至少交换 3 个数据报文，也就是所谓的 3 次握手事件，如图 2.12 所示为 3 次握手建立连接的示意图。

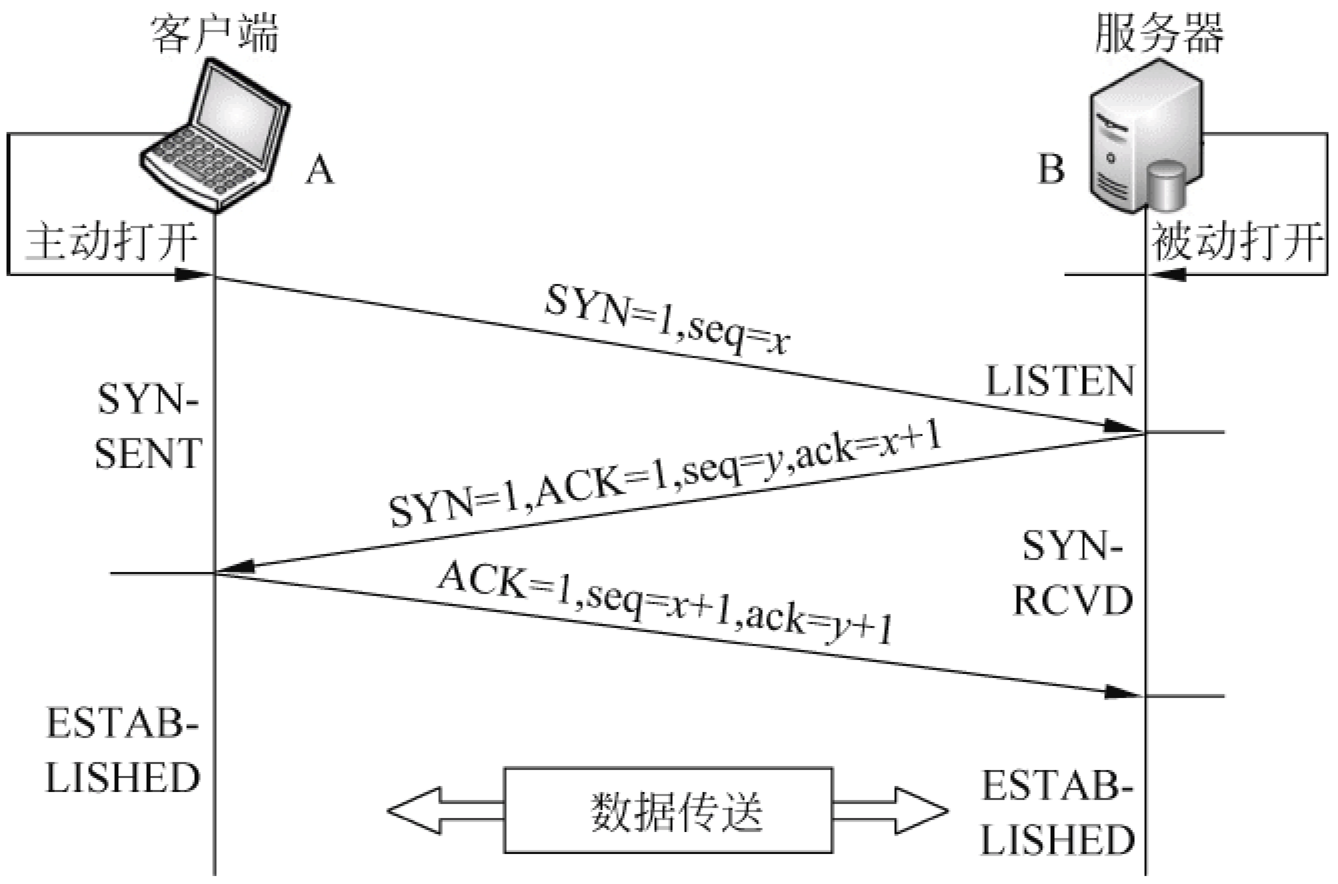


图 2.12　3 次握手建立 TCP 连接

（1）首先服务器被动打开，准备接收外来的连接。

（2）客户端主动打开发送同步序号（SYN＝1）以及初始序号（seq＝x），并在这个不携带数据的报文中要指明想要连接的服务器的端口。

（3）服务器确认收到客户端发来的 SYN 信号，并在发送确认信息（ACK＝1）的同时也发送同步序号（SYN＝1）信号，包含服务器在同一连接中要发送的初始序号（seq＝y）。ACK 确认信号中的序号（ack＝x＋1）为客户端的初始序号值 x 加 1 来对客户端的连接进行确认，同样也为自己选择一个初始序号（seq＝y）。

（4）最后客户端再回传确认信号（ACK＝1）对服务器发来的 SYN 信号进行确认，建立连接，这时 ack＝y＋1，seq＝x＋1。

2）TCP 通信传输数据过程

（1）客户端在发送数据前先将数据存放在发送缓冲区内，再用 TCP 建立一个段，段内包含指明该段序列号的报头。

（2）该数据段被封装成 IP 数据报，传输给服务器。

（3）当 TCP 发出一个段后，就启动一个定时器，等待目的端确认收到这个段。此时该段仍保留在发送缓冲区内。如果不能及时收到一个确认，将重发这个段。

（4）服务器收到客户端的数据后，将发送一个确认，指明已经接收的字节数目。通常这个确认在推迟几分之一秒后发送。该确认中包含有当前窗口的尺寸。

（5）客户端收到确认后，将确认的段从发送缓冲区删除，并重传未得到确认的段。确认将向前移动窗口以发送新的段。

3）TCP 通信的可靠性概述

（1）TCP 实体把应用程序划分为合适的数据块，加上 TCP 报文头，生成数据段。

（2）当 TCP 实体发出数据段后，立即启动计时器，如果源设备在计时器清零后仍然没有收到目的设备的确认报文，重发数据段。

（3）当对端 TCP 实体收到数据，发回一个确认。

(4) TCP 包含一个端到端的校验和字段,检测数据传输过程的任何变化。如果目的设备收到的数据校验和计算结果有误,TCP 将丢弃数据段,源设备在前面所述的计时器清零后重发数据段。

(5) 由于 TCP 数据承载在 IP 数据包内,而 IP 提供了无连接的、不可靠的服务,数据包有可能会失序。TCP 提供了重新排序机制,目的设备将收到的数据重新排序,交给应用程序。

(6) TCP 提供流量控制。TCP 连接的每一端都有缓冲窗口。目的设备只允许源设备发送自己可以接收的数据,防止缓冲区溢出。

(7) TCP 支持全双工数据传输。

4) TCP 断开连接

在 TCP 断开连接时,需要进行 4 个数据包的交换,即 4 次握手,如图 2.13 所示。

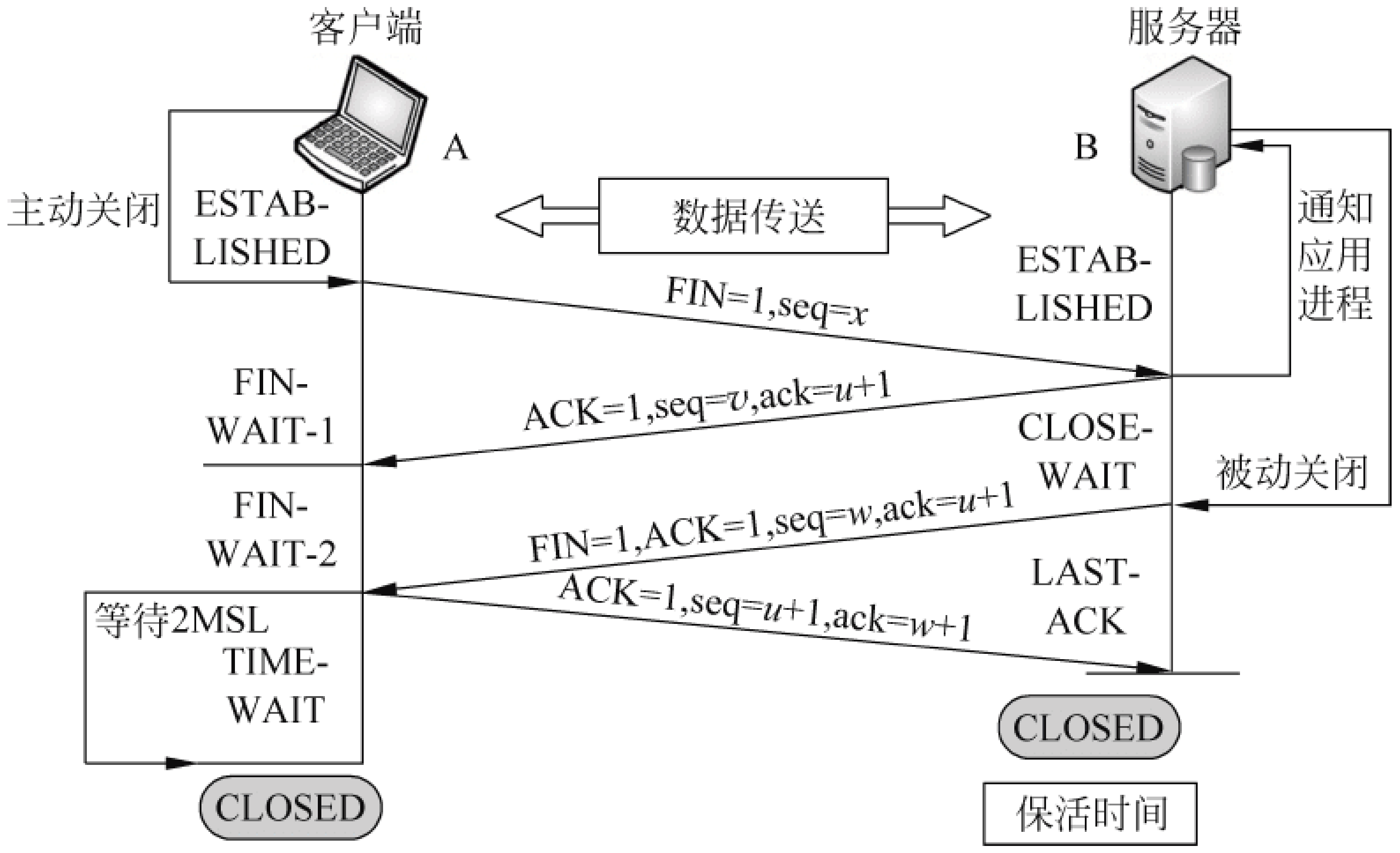


图 2.13 断开 TCP 连接

(1) 在数据传输结束后客户端主动关闭,并向服务器端发送结束信号(FIN=1)。

(2) 服务器在接收到 FIN 后,向响应需要结束的应用程序传送文件结束符,通知应用进程进行被动关闭,并向发来信号的客户端返回一个确认信号ACK,此时客户端进入等待状态等待服务器关闭相应进程并传回确认信号。

(3) 等待一段时间,收到文件结束符的服务器程序关闭后,向对应客户端发送一个结束符号 FIN 以及一个确认关闭信号 ACK。

(4) 在客户端收到该信号后会进入等待状态(等待时间 2s,之后会进行关闭,防止服务器发来的确认报文丢失而造成无法释放连接的情况),并回复一个确认信息 ACK,当服务器接收到这个信息后会断开连接(此时会存在一个保活时间,即使客户端突然死机了,服务器也会在保活时间到了之后发送一个探测消息来决定是否要断开连接)。

【例 3】 主机 A 向主机 B 连续发送了两个 TCP 报文段,其序号分别是 60 和 100,试问:

(1) 第 1 个报文段携带了多少字节的数据?

答: 因为第 1 个报文序号 60~99,所以共 40 字节。

(2) 指明第 1 个报文中的窗口值是多少?

答: 所谓窗口值就是指对端准备接收的字节数目,当然也应该为 40 字节。

(3) 在主机 B 收到第 1 个报文后发回的确认报文中,确认号应该是多少?

答: 确认号应该是要求对端将要发送的下一个报文的第 1 个序号,即 100。

(4) 如果 B 收到第 2 个报文段后,发回的确认中的确认号是 180,试问 A 发送的第 2 个报文段中的数据有多少个字节?序号为多少?

答: 发送的报文应该是 180−100=80 字节,即序号为 100~179。

(5) 如果 A 发送的第 1 个报文段丢失了,但第 2 个报文段到达了 B。B 在第 2 个报文段到达后向 A 发送确认,试问这个确认应为多少?

答: 只有重新再发,即 60。

【例 4】 设 TCP 使用的最大窗口为 65 000 字节,而传输信道不产生差错,带宽也不受限制。若报文段的平均往返时间为 20ms,问所能得到的最大吞量是多少?

答: 最大窗口可以理解为最大报文段长度(MSS),即数据字段加上 TCP 首部才等于整个的 TCP 报文段。在发送时延可忽略的情况下,每 20ms 可发送 65 000×8=520 000 位

最大数据率=(520 000 位)/(20ms)=26Mbps。

2.4.3 UDP 和 SCTP

1. UDP

UDP 主要是面向请求/应答式的交易型应用,它应用于那些对可靠性要求不高,但要求网络的延迟较小的场合,如语音和视频数据的传送。UDP 尽管不保证数据报能到达目的地,但由于网络技术的快速进步,UDP 的用途显得尤为重要,如第四代移动通信(4G)就是采用 UDP 来传输数据业务。UDP 报文格式如图 2.14 所示。

0　　　　　　　　15	16　　　　　　　　31
16位源端口号	16位目的端口号
16位UDP长度	16位UDP校验和
数据(如果有)	

图 2.14　UDP 报文格式

(1) 16 位源端口号:指明了发送端应用程序所在的端口。

(2) 16 位目的端口号:指明了接收端应用程序所在的端口。

(3) 16 位 UDP 长度:指明了数据报的长度,最小 8 字节,即 UDP 报头长度。

(4) 16 位 UDP 校验和:是覆盖了 UDP 首部和 UDP 数据的校验和。

UDP 的优势就是只有较少的控制选项,在数据传输过程中,延迟较小,数据传输效率较高。UDP 最适合对可靠性要求不高的应用程序,如 DNS、TFTP、SNMP 等。

2. SCTP

SCTP 是在 IP 网络使用的一种可靠的通用传输层协议,它通过借鉴 UDP 的优点解决了 TCP 的某些局限,SCTP 已经逐渐发展成为一种通用的传输层协议,它除了具有 TCP 同样的功能之外,还具有更灵活的数据报格式,能更好地扩展以满足某些应用的需求,其主要特征如下。

(1) 内建多地址主机支持:SCTP 中的一对连接称为偶联(association),偶联两端的主机节点(endpoint)可以有多个网络地址,如 4G 控制面的信令的传送,就是采用 SCTP 多条

偶联通路进行数据传输。偶联就是两个端点通过 SCTP 协议规定的 4 次握手建立起来的逻辑通路(path),有的文献也称关联。

(2) 保留应用层消息边界:SCTP 保留上层数据信息的边界,上层数据信息称为“消息”,传输的基本单位为有意义的数据段。

(3) 单个偶联(association)多流机制:SCTP 允许用户在每个偶联中定义子流,数据在子流内按序传输。

SCTP 公共分组头和数据块格式如图 2.15 所示。

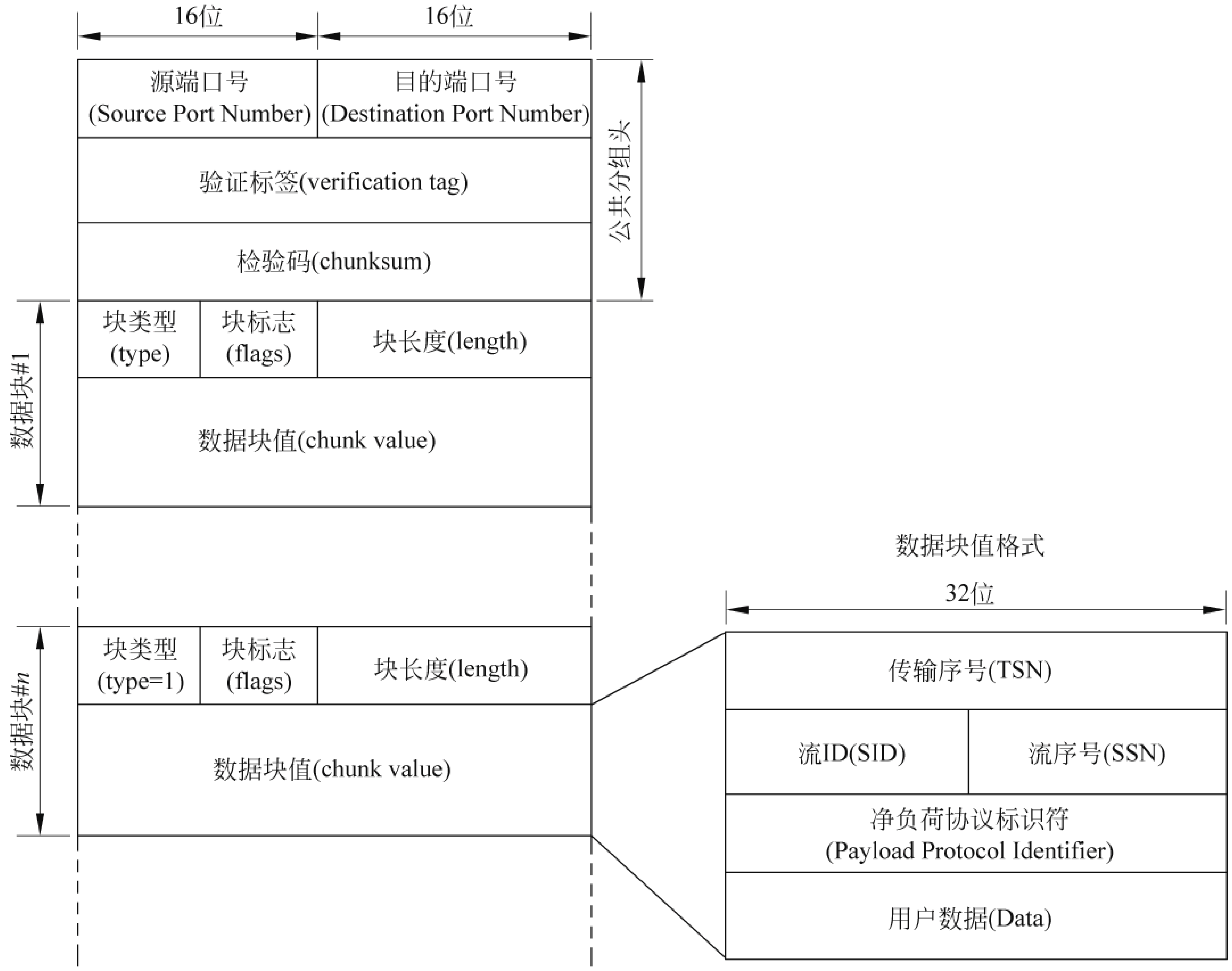


图 2.15　SCTP 公共分组头和数据块格式

(1) 公共分组头格式。

源端口号(source port number):16 位,为发送端 SCTP 端口。

目的端口号(destination port number):16 位,为接收端 SCTP 端口。

验证标签(verification tag):32 位,是偶联建立时,本端端点为这个偶联随机生成的一个标识。

检验码(chunksum):32 位,ADLET-32 算法生成的一个校验码。

(2) 数据字段(chunk)格式。

块类型(chunk type)8 位,指块值中的消息类型。

块标志(chunk flags)8 位,其用法由块类型决定。

块长度(chunk length)16 位,表示数据块的总长度,必须为 4 倍字节的整数倍。

数据块值(chunk value)：若块类型 Type=1 时，其值为用户数据，格式内容如下：

① 传输序号(Transmission Sequence Number，TSN)：32 位，表示该数据块的序号。

② 流 ID(stream ID)：16 位表示用户数据属于的流。

③ 流序号(Stream Sequence Number，SSN)16 位，表示流中的用户数据序号。

④ 净负荷协议标识符(payload protocol identifier)：32 位，是上层协议给定的一个标识符，表示某个应用。

⑤ 用户数据(data)：即为上层数据。

2.5 应用层

应用层是 OSI 参考模型最接近用户的一层，应用层支持应用程序。应用层负责处理特定的应用程序细节。应用层显示接收到的信息，把用户的数据发送到低层，为应用软件提供网络接口。应用层包括所有的高层协议，表 2.6 给出了常用协议号及其对应端口号。

表 2.6 常用协议号及其对应端口号

应　用	协议	对应 TCP 报文的端口号	对应 UDP 报文的端口号	对应 IP 报文的协议号	对应 MAC 帧的协议号
万维网	HTTP	80(WWW)			
文件传送	FTP	20(数据)/21(控制)			
远程终端接入	Telnet	23			
电子邮件	SMTP	25			
边界网关	BGP	179			
域名转换	DNS		53		
引导程序	BootP		67/68(client 收/发)		
动态主机分配	DHCP		67/68(BootP 改进)		
简单文件传送	TFTP		69		
网络管理	SNMP		161/162		
内部路由选择	RIP		520		
传输控制	TCP			6	
数据报传输	UDP			17	
内部路由选择	OSPF			88、89	
报文控制	ICMP			1	
网络互连	IP				0

早期的应用层有远程登录协议(Telnet)、文件传输协议(FTP)和简单邮件传输协议(SMTP)等。后来出现了一些新的应用层协议：如用于将网络中的主机的名字地址映射成网络地址的域名服务(DNS)；网络新闻传输协议(NNTP)和用于从 WWW 上读取页面信息的超文本传输协议(HTTP)等。下面简单介绍常用的几种应用层协议。

(1) 文件传输。文件传输协议(File Transfer Protocol，FTP)是用于文件传输的 Internet 标准。FTP 支持一些文本文件(例如 ASCII、二进制等)和面向字节流的文件结构。FTP 使用传输层协议 TCP 在支持 FTP 的终端系统间执行文件传输，因此，FTP 被认为提供了可靠的面向连接的服务，适合于远距离、可靠性较差的线路上的文件传输。

超文本文件传输协议(Trivial File Transfer Protocol,TFTP)也是用于文件传输,但 TFTP 使用 UDP 提供服务,被认为是不可靠的、无连接的。TFTP 通常用于可靠的局域网内部的文件传输。

(2) 邮件服务。简单邮件传输协议(Simple Mail Transfer Protocol,SMTP)支持文本邮件的 Internet 传输。

邮局协议第 3 版本(Post Office Protocol 3,POP3)是一个流行的 Internet 邮件标准。

(3) 网络管理。简单网络管理协议(Simple Network Management Protocol,SNMP)负责网络设备监控和维护,支持安全管理、性能管理等。

Telnet 是客户机使用的与远端服务器建立连接的标准终端仿真协议。

ping 命令是一个诊断网络设备是否正确连接的有效工具。

tracert 命令和 ping 命令类似,tracert 命令可以显示数据包经过的每一台网络设备信息,是一个很好的诊断命令。

(4) 网络服务。HTTP 协议支持万维网(World Wide Web,WWW)和内部网信息交互,支持包括视频在内的多种文件类型。HTTP 是当今流行的 Internet 标准。

域名系统(Domain Name System,DNS)把网络节点的易于记忆的名字转化为网络地址。

Windows Internet 命名服务器(Windows Internet Name Server,WINS),此服务器可以将 NetBIOS 名称注册并解析为网络上使用的 IP 地址。

引导协议(Bootstrap Protocol,BootP)是使用传输层 UDP 协议动态获得 IP 地址的协议。

【例 5】 某 PC 通过局域网连接到了 Internet,简要说明该 PC 在打开 Internet 中的一个网页时可能会用到哪些通信协议,并参考表 2.6 说明各协议的用途及其协议号的封装过程。

答:

(1) 每层用到的通信协议:①应用层:DNS 协议号为 53、HTTP 协议号为 80;②传输层:TCP 协议号为 5、UDP 协议号为 17;③网络层:IP 协议号为 0;④数据链路层以下:以太网(802.3 协议)。

(2) DNS 的请求进程。

PC 在要想访问 Internet 中的某一个网页,必须要首先知道这个网页所在服务器的 IP 地址,就要调用 DNS,通过域名服务器将网页的统一资源定位符(Uniform Resource Locator,URL),转换成网站的 IP 地址。URL 是 WWW 的统一资源定位标志,由资源类型、存放资源的主机域名、资源文件名组成。UDP 用于传输 DNS 请求和响应。其请求进程封装过程如下:①应用层将 DNS 的请求进程交给传输层 UDP,在 UDP 报文中要封装的端口号就是 DNS 协议号即 53;②传输层将 UDP 报文交给网络层的 IP,在 IP 报文中要封装的协议号为 UDP 的 17;③网络层的 IP 报文要交给以太网,以太网的 MAC 帧中就要封装 IP 的协议号,其类型字段为 0800,后面 00 就表示它的上一层协议号为 0,即 IP 协议号;④物理层通过以太网接口,将数字数据信息送到网络的连接设备,如交换机。

(3) HTTP 用于请求和传输网页:①应用层将 HTTP 用于请求和传输网页的进程交给传输层 TCP,在 TCP 报文中要封装的端口号就是 HTTP 协议号即 80;②TCP 要启动 3

次握手，同对端进行建立连接。如第1次握手就是主动打开发送同步序号SYN以及初始序号，并指明需要连接的服务器的端口为80；③传输层将每一次建立连接以及传输数据的TCP报文，都要交给网络层的IP协议，在IP报文中要封装的TCP协议号为6，就表明传输的上层协议是TCP报文；④网络层的IP报文要下交给以太网，以太网的MAC帧中就要分装IP的协议号，然后交给物理层，通过以太网接口，将数字数据信息送到网络的连接设备。

习题

一、单项选择题

1. (　　)协议可以根据已知的IP地址确定MAC地址。
   A. ARP　　B. RARP　　C. RAP　　D. ICMP
2. 在计算机网络的层次结构中，第 N 层通过它的服务访问点向(　　)提供服务。
   A. 第 N 层　　B. 第 $N+1$ 层　　C. 第 $N-1$ 层　　D. 最顶层
3. 在计算机网络中，(　　)负责实现分组的路由选择功能。
   A. 物理层　　B. 数据链路层　　C. 网络层　　D. 传输层
4. 计算机网络中各节点之间传输方式采用(　　)。
   A. 串行方式　　B. 并行方式　　C. 连续方式　　D. 分散方式
5. UDP提供面向(　　)的传输服务。
   A. 端口　　B. 地址　　C. 连接　　D. 无连接
6. 在下列协议中，全部属于网络层协议的是(　　)。
   A. IP、TCP、UDP　　B. ARP、IP和UDP
   C. FTP、DSP、TCP　　D. ICMP、ARP、IP、RARP
7. 把IP地址转换为以太网MAC地址的协议是(　　)。
   A. DNS　　B. ARP　　C. RARP　　D. DHCP

二、多项选择题

1. TCP/IP协议族中的网络接口层对应于OSI参考模型的哪几层？(　　)
   A. 网络层　　B. 数据链路层　　C. 会话层　　D. 物理层
2. IP数据报的首部中，与分段重装有关的字段是(　　)。
   A. 偏移字段　　B. 标识字段　　C. 标志字段　　D. 片偏移字段
3. 下列关于IP地址的描述哪些是正确的？(　　)
   A. Internet上的每个接口必须有一个唯一的IP地址
   B. Internet上的每个主机必须有一个唯一的IP地址
   C. IP地址可以分为5类
   D. 32位全为“1”的IP地址表示在整个Internet中广播
4. 在下列协议中，哪些是IP层协议？(　　)
   A. IP　　B. X.21　　C. IGMP　　D. FTP

三、问答题

1. IP地址分几类？有哪些是特殊的IP地址？
2. 192.168.3.0/24表示一个网络地址还是主机地址？

3. 192.168.3.0/28 通常可以划分多少个子网？每个子网中有多少个主机地址？为什么？

4. 为什么说 IP 层是一个不可靠的通信？

5. TCP、UDP 和 SCTP 有什么不同？

6. 两句话："过河前要先建桥，过河以后要拆桥"；"摸着石头过河"。哪一句适合于无连接？哪一句适合于面向连接？有人说"IP 网本身就是一个无连接的网络，所以 TCP 也不是真正意义上的面相连接，只是一种受控状态下的传输"，你认为对吗？为什么？

四、讨述题

1. OSI 参考模型和 TCP/IP 参考模型有什么区别？分别有哪些特点？

2. IP 首部中哪些字段与 IP 数据报的分段重装有关？一个 IP 数据报在什么情况下需要分段？各分段在何处进行重装？

3. 一个 TCP 数据报文长度为 4100 字节(包括固定首部长度)。现要经过一个网络层传输，但此网络层能够传送的最大数据长度为 1500 字节。试问应该划分为几个短的数据报片？各数据报片的数据长度、片偏移字段和 MF 标志为何值？

第 3 章 Internet

Internet 有一个专用名词“互联网”，又称网际网络，或译因特网，指当前全球最大的、开放的、由众多网络相互连接而成的特定的计算机网络，它采用 TCP/IP 协议簇。Internet 是建立在不同硬件结构网络、MAC 地址之上的虚拟网，用统一的 IP 地址实现网络的互联，本章主要介绍 Internet 设备、Internet 子网划分、应用层协议和 Internet 接入等技术。

3.1 Internet 设备

3.1.1 网络设备概念模型

计算机网络，就是将分散的具有独立功能的多台计算机互相连接在一起，在网络操作系统、管理软件及通信协议的管理和协调下，实现资源共享和信息传递的数据通信。网络互连设备工作在网络模型的不同层，并根据不同层的特点完成各自不同的功能，图 3.1 给出了网络设备与 OSI 七层模型协议的对应关系概念模型，当然计算机网络模型的应用层包含了最上面的三层，下节将围绕这个模型介绍具体设备。

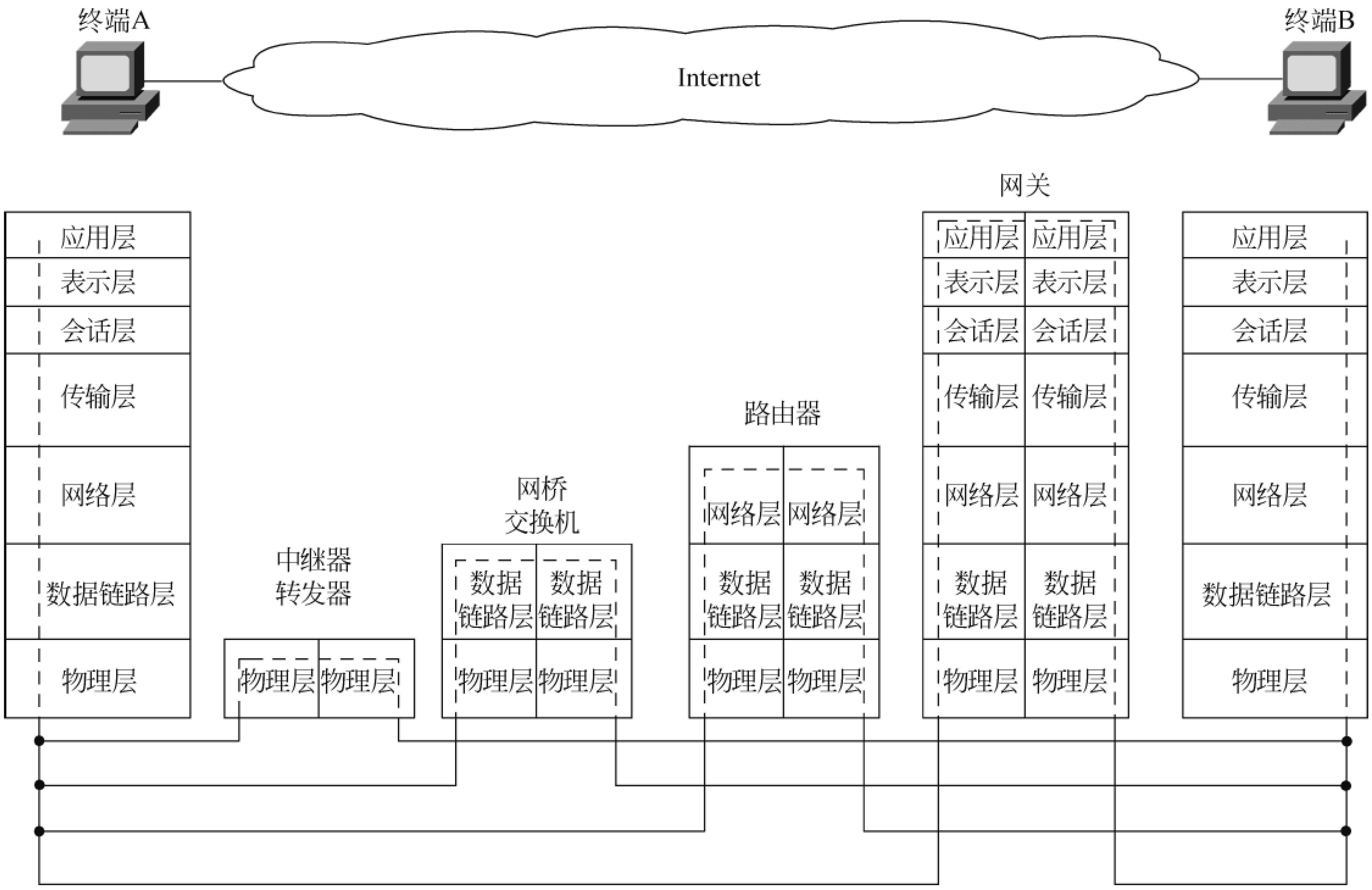


图 3.1　网络设备概念模型

3.1.2 网络设备

1. 转发器(repeater)

转发器也称中继器,指物理层互连设备。物理层的连接,只是信息的复制,以增加两端的长度。因此,用中继器连接的两个网段仍共同处于一个冲突域。

在实际组网时,每种传输介质其传输距离都是有限的,例如,粗同轴电缆每一网段的最大距离为 500m,细同轴电缆为 180m,双绞线为 100m,超过这些距离,就需要利用转发器来扩展距离。转发器的功能就是将经过衰减而变得不健全的信号,经过整理、放大后,再继续传送。

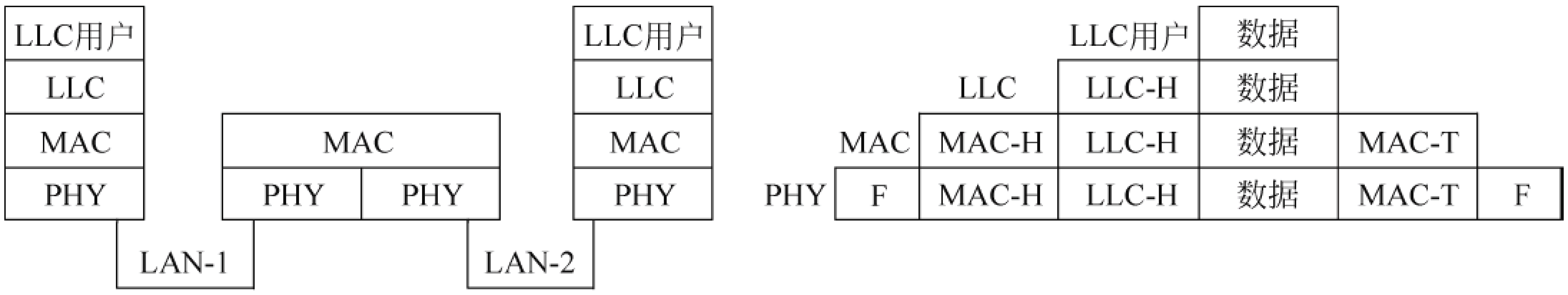
集线器(hub)也属于中继器,可以看成是一种多接口的中继器,是总线型共享带宽式的,其每个接口的带宽就是总带宽。

交换式集线器(也称局域网交换机)属于二层设备,每一接口都有其专用的带宽,如 100Mbps 的交换式集线器,指每个接口都有 100Mbps 的带宽。

交换式集线器和集线器普遍采用自适应技术,自动适应 100Mbps 和 10Mbps 速率。这类交换式集线器和集线器按照以下顺序适应工作速率:100Mbps 全双工、100Mbps 半双工、10Mbps 全双工、10Mbps 半双工。在不需用户参与设定的情况下,自动以最高速率连接。

2. 网桥(bridge)

网桥也称桥接器,指数据链路层互连设备,链路层的帧中继。图 3.2 为网桥的连接及其帧结构。用于连接一个局域网的两个网段 LAN-1 和 LAN-2,即两个冲突域,每个网段里的机器是一个冲突域,但网桥连接的两个网段同处于一个广播域,仍属于一个局域网 LAN。



(a) 通过网桥连接的局域网 (b) 对应层结构

图 3.2 网桥连接及其帧结构

图 3.2(a)是当 MAC 帧目标地址和源地址属于不同的 LAN 网段时,该帧被网桥捕获、缓存和转发,两个网段之间对于对等逻辑链路控制子层(LLC)实体之间有对话,但是网桥只是打开 MAC 帧,用以识别 MAC 地址,而不需要知道 LLC 地址。图 3.2(b)给出的是针对各个链路子层以及物理层形成的报文格式,其中,数据表示高层进入的数据报文,如 IP 报文等;LLC-H 表示 LLC 帧头;MAC-H 表示 MAC 帧头,MAC-T 表示 MAC 帧尾;F 为物理层增加的隔离位。

网桥具有信号过滤的功能,对每个帧进行分析,根据信宿物理地址(如 MAC 地址)来决定数据的去向。网桥具有传输高层协议的透明性,在数据链路层上操作,无须检查高层的信息。

传统的网桥只有两个接口,而多接口的网桥则称为交换机。

3. **交换机**(switch)

通常说的交换机就是指二层交换机,也称为交换式集线器或以太网交换机,它的出现是为了解决连接在集线器上的所有主机共享可用带宽的缺陷,它通过为需要通信的两台主机直接建立专用的通信信道来增加可用带宽。交换机没有过滤广播通信的功能,所有接口所连接的主机共同构成一个广播域。

数据链路层以上,完成报文交换的互连设备,目前有 ATM、以太网交换机或交换机(二层)、三层交换机,四层交换机、应用层交换机等。也可以将路由器和二层交换机的功能集成到一起,称作路由交换机或交换路由器。多层交换机可提供路由转换、多协议转换、包交换等功能。目前,有些以太网交换机也具有了一定的路由功能,是一种二层、三层合用的设备。

4. **路由器**(router)

路由器指网络层互连设备。网络层上实现网络的互连,是一种智能型节点设备,具有连接、地址判断、路由选择、数据处理和网络管理功能,并对数据报进行检测,决定发送方向。因为它处于网络层,一方面能够跨越不同的物理网络类型,如 DDN、FDDI、ATM 等,另一方面将整个互连网络分割成逻辑上相对独立的网络单位,使网络具有一定的逻辑结构。路由器有能力过滤广播消息。实际上,除非用做特殊配置,否则路由器从不转发广播类型的数据报文。因此,路由器的每个接口所连接的网络都独自构成一个广播域。

路由器是网络中进行网间连接的关键设备,它的基本功能是把数据报传送到正确的网络中去,具体包括:数据报的转发、子网隔离、维护路由表、差错处理及简单的拥塞控制、对数据报的过滤和记忆,并与其他路由器交换路由信息等功能。

5. **网关**(gateway)

网关又称网间连接器、协议转换器。指网络层以上互连设备。可连接不同工作协议的主机设备,通过对不同协议的转换,实现网络间的互连,通常用于异型网的连接,要求顶层协议相同。

网关是一个翻译器,是一种充当转换重任的计算机系统或设备。使用在不同的通信协议、数据格式或语言,甚至体系结构完全不同的两种系统或网络之间。网关通常可以是安装在路由器内部的软件,结合路由器上的协议,实现协议的转换,将信息从一个网络传输到另一个网络;网关也可能是在一个服务器或 PC 上安装相关的网关软件,实现网关功能。

6. **四层交换机**(layer-4 switch)

四层交换机也称会话交换机,是基于“数据流”的概念来进行数据报文的寻径转发的,可以实现“一次路由,多次交换”。

四层交换的特点为:可通过检查端口号,识别不同报文的应用类型,从而根据应用对数据流进行分类;可根据数据流应用类型,提供 QoS 和流量统计;网络中传输的数据可以认为是在特定的时间内,由特定的目的和源之间的数据流组成的;可依据设备数据流的信息对数据报文实现交换;可方便网络管理者根据数据流的应用类型,定义不同的优先级和 QoS。例如,可以为视频用户分配高优先权、带宽等。

每台四层交换机都保存一个与被选择的服务器相匹配的源 IP 地址,以及与源 TCP 端口相关联的连接表,然后四层交换机向这台服务器转发连接请求。所有后续包在客户机与服务器之间重新映射和转发,直到交换机发现会话为止。四层交换中数据报文的传输不仅

依据 MAC 地址或源/目标 IP 地址，还要依据 TCP/UDP 的端口地址。端口地址代表了不同的业务协议，所以第四层交换不仅进行了物理上的交换，还包括了业务上的交换；第四层交换的交换域是由源端和终端 IP 地址、TCP 或 UDP 端口共同决定的，因此，也称会话交换机。

7．服务器（server）

大多数服务器是网络的核心（当然对等网也可以没有服务器）。作为普通的办公、教学等应用，服务器可以采用一般配置较高的普通计算机；应用要求高的地方，如证券所等，则采用专用的服务器。专用服务器比普通计算机具有更好的安全性和可靠性，更加注重系统的 I/O 吞吐能力，一般采用双电源、热拔插、SCSI RAID 硬盘等技术。

8．网络适配器（network adapter）

网络适配器指网卡等，为主链路以下设备，它的主要作用是将计算机数据转换为能够通过介质传输的信号。当网络适配器传输数据时，它首先接收来自计算机的数据，对数据附加网卡地址等报头，然后将数据转换为可通过传输介质发送的信号。

9．防火墙（fire wall）

在互联网的子网（或专用网）与互连网之间设置的安全隔离设施，可提供接入控制，干预两网之间各种消息的传递等，达到网络和信息安全的效果。

10．IP 交换（IP switching）

IP 交换即 IP 交换机。IP 交换机可以由 ATM 交换机和 IP 交换控制器组成。

11．IP 电话（IP phone）

在 IP 网上提供的具有一定服务质量的语音业务的终端设备。

12．（IP 电话）网守[（IP Phone）gatekeeper]

网守也称关守，是在 IP 电话网上提供地址解析和接入认证的设备。

13．网络终端（network terminal）

曾经的概念强调是专用于网络计算环境下的终端设备，现在的概念概括为所有终端设备，通常都具有由高层到低层的完备功能，如 PC、智能手机、平板电脑等能上网的设备，它们通过网络获取资源，应用软件和数据也都存放在服务器上，通常也称为机器、主机等。

14．七层交换机（layer-7 switch）

七层交换机或称高层智能交换，第七层交换技术可以定义为数据报文的传送不仅仅依据 MAC 地址、IP 地址以及 TCP/UDP 端口，而且可以根据内容进行传送。七层交换是以进程和内容级别为主的交换。高层由于和应用相关，这时候的交换就有了智能性，交换机具有了区别各种高层应用和识别内容的能力。这时的交换机不仅能根据数据报文的 IP 地址、端口地址来传送数据，而且还能打开数据报文，进入数据报文内部，根据报文中的信息做出负载均衡、内容识别等判断。目前关于第七层交换功能还没有具体的标准。

15．Web 交换机

高层交换技术的一个典型应用就是 Web 交换机。在目前 Internet 网站上的信息量和访问急剧增长的前提下，怎样使每个用户都可以得到 QoS 的保证是一个越来越重要的问题。所以必须提供一种处于中心地位的交换机，即 Web 交换机，来组织数据中心交换。Web 交换机的基本功能是：组织数据中心，提供对外一致服务界面，管理数据的流向与路由，负载均衡，提供 QoS 和 CoS（服务等级）以及请求会话定向。为了实现上述功能，基于内

容 Web 交换机必须检查 4～7 层的协议字段来获取信息，以处理数据流的管理和定向。目前，Web 交换机主要设计方案有 3 种：集中式 CPU 模式、分布式处理系统和二级混合模式。

3.2 IP 地址

目前，网络地址规划主要采用子网编址、无子网编址、VLSM 和 CIDR 等技术。

3.2.1 子网划分

1. IP 地址分类

前一章已经讲过，IP 地址分为 5 类，以下是 IP 地址分类。

(1) A 类 1.0.0.0～126.0.0.0 有效，0.0.0.0 和 127.0.0.0 保留。

(2) B 类 128.1.0.0～191.254.0.0 有效，128.0.0.0 和 191.255.0.0 保留。

(3) C 类 192.0.1.0～223.255.254.0 有效，192.0.0.0 和 223.255.255.0 保留。

(4) D 类 224.0.0.0～239.255.255.255，用于多点广播。

(5) E 类 240.0.0.0～255.255.255.254 保留，255.255.255.255 用于广播。

当建立内部网的时候，可使用私有网络地址，以下是 3 类私有网络地址分配。

(1) A 类私有网络地址：10.1.1.1～10.254.254.254。

(2) B 类私有网络地址：172.16.1.1～172.31.254.254。

(3) C 类私有网络地址：192.168.1.1～192.168.254.254。

公有地址是除私有地址以外的其他 IP 地址，它们可以在 Internet 上进行路由，也称为合法的 IP 地址。

2. 无子网编址与子网编址

无子网编址是指使用自然掩码，不对网段进行细分。对于没有子网的 IP 地址组织，外部将该组织看作单一网络，不需要知道内部结构。例如，如图 3.3(a)所示的 B 类网段为 172.12.0.0，采用掩码为 255.255.0.0。所有到地址 172.12.x.x 的路由被认为同一方向，不考虑地址的第三和第四个 8 位组，这种方案的好处是减少路由表的项目。

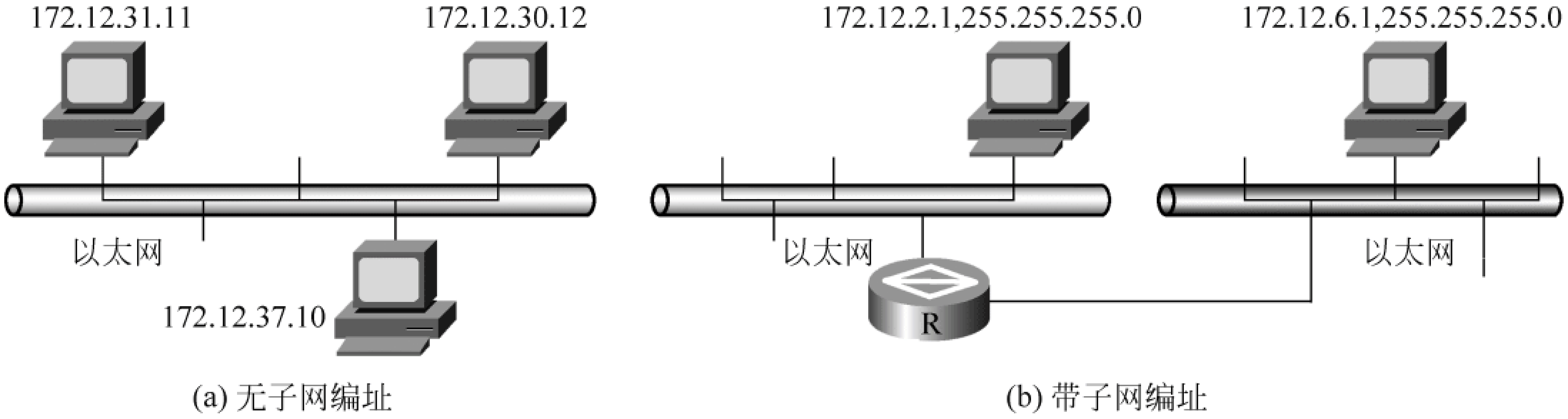


图 3.3　B 类网段划子网与不划子网的情况

采用无子网编址这种规划方案，不足之处就是使得网络内所有主机都能收到在该网络内的广播，增加流量，降低性能，也不利于管理。如果一个网络要容纳上千个主机时，要同时管理这么多主机是相当困难的，这就需要一种方法将这种网络分为不同的网段，也就是子网，然后按照各个子网段进行管理。从地址分配的角度来看，子网是网段地址的扩充。

网络设备使用子网掩码(subnet masking)决定 IP 地址中哪部分为网络部分,哪部分为主机部分。子网掩码使用与 IP 地址一样的格式。子网掩码的网络部分和子网部分全都是 1,主机部分全都是 0。

如图 3.3(b)采用 B 类网段,将网络 172.12.0.0 分为两个子网段:172.12.2.0、172.12.6.0。如果公司的商务部使用 172.12.2.0 子网段,公司的制造部使用 172.12.6.0 子网段,这样可使路由器根据目的子网地址进行路由,从而限制一个子网的广播报文发送到其他网段。

3. 子网掩码及子网划分

IP 划分子网时,要为每台设备指定子网掩码,掩码决定了网络地址的长度,具有相同子网掩码的设备就处于同一个子网内。掩码用于识别 IP 地址中的网络地址位数,IP 地址(ip-address)和掩码(mask)相与即得到网络地址。通过使用可变长的子网掩码,可以合理调度 IP 地址,有效组网,充分利用有效的 IP 地址空间。

为了方便说明问题,图 3.4 给出了 B 类地址子网及子网掩码的划分。掩码最前面部分连续二进制“1”的个数为 22 位,表示网络地址为 22 位,则掩码为 255.255.252.0,也就是子网地址(subnet-id)占用了 6 位主机地址(host-id)。如图 3.5 所示为十进制和二进制的地址转换表。

B类地址:	10NNNNNN	net-id		host-id	host-id
B类带子网地址:	10	net-id		subnet-id	host-id
子网掩码:	11111111	11111111		11111100	00000000

图 3.4 B 类地址子网及子网掩码的划分

128	64	32	16	8	4	2	1	
0	0	0	0	0	0	0	0	=0
1	0	0	0	0	0	0	0	=128
1	1	0	0	0	0	0	0	=192
1	1	1	0	0	0	0	0	=224
1	1	1	1	0	0	0	0	=240
1	1	1	1	1	0	0	0	=248
1	1	1	1	1	1	0	0	=252
1	1	1	1	1	1	1	0	=254
1	1	1	1	1	1	1	1	=255

图 3.5 十进制和二进制数的转换

表 3.1 为各种类型的网络地址范围及对应的掩码,A 类网络的子网掩码为 255.0.0.0,B 类网络的子网掩码为 255.255.0.0,C 类网络子网掩码为 255.255.255.0。利用子网,网络地址的使用会更有效,对外它们仍处在一个网络内,对内则是分为不同的子网。通过使用可变长的子网掩码可以让位于不同接口的同一网络编号的网络使用不同的掩码,这样可以充分利用有效的 IP 地址空间。

表 3.1 各种类型的网络地址范围及对应的掩码

类型	网络(N).主机(H)	网络地址范围	标准二进制掩码
A	N. H. H. H	1～126	1111 1111 0000 0000 0000 0000 0000 0000
B	N. N. H. H	128～191	1111 1111 1111 1111 0000 0000 00000000
C	N. N. N. H	192～223	1111 1111 1111 1111 1111 1111 00000000

表 3.2、表 3.3 给出了 B、C 类型子网位个数对应的网络及主机数，表中的子网数是不含全 0 全 1，要注意的是现在大部分子网基本上是都支持子网全 0、全 1 配置的，一定要区别对待。主机数是必须要减去全 0 全 1 的，必要时还要再减去 1 个"网关地址"。

表 3.2 B 类地址子网表

子网位数	子网掩码	子网数	每一子网主机数
2	255.255.192.0	2	16 382
3	255.255.224.0	6	8190
4	255.255.240.0	14	4094
5	255.255.248.0	30	2046
6	255.255.252.0	62	1022
7	255.255.254.0	126	510
8	255.255.255.0	254	254
9	255.255.255.128	510	126
10	255.255.255.192	1 022	62
11	255.255.255.224	2 046	30
12	255.255.255.240	4 094	14
13	255.255.255.248	8 190	6
14	255.255.255.252	16 382	2

表 3.3 C 类子网地址表

子网位数	子网掩码	子网数	每一子网主机数
2	255.255.255.192	2	62
3	255.255.255.224	6	30
4	255.255.255.240	14	14
5	255.255.255.248	30	6
6	255.255.255.252	62	2

4. 子网划分举例

【例 1】 某单位分配到一个 B 类 IP 地址，其网段地址为 129.253.0.0，分别回答以下问题。

(1) 如果选用掩码为 255.255.240.0，可以分为多少个子网，每个子网含有多少个主机？

答：该子网及主机分配如图 3.6 所示，已知子网掩码为 255.255.240.0 的 B 类地址，即子网个数取 4 位时，就可以有 $2^4=16$ 个子网，包含全 0 和全 1 子网；每个子网中含有 $2^{12}-2=4094$ 个主机，当然也可以通过查表 3.2 得到这个数。

(2) 该单位共有 4000 台主机，如主机平均分布在 16 个不同地点，如果子网掩码为

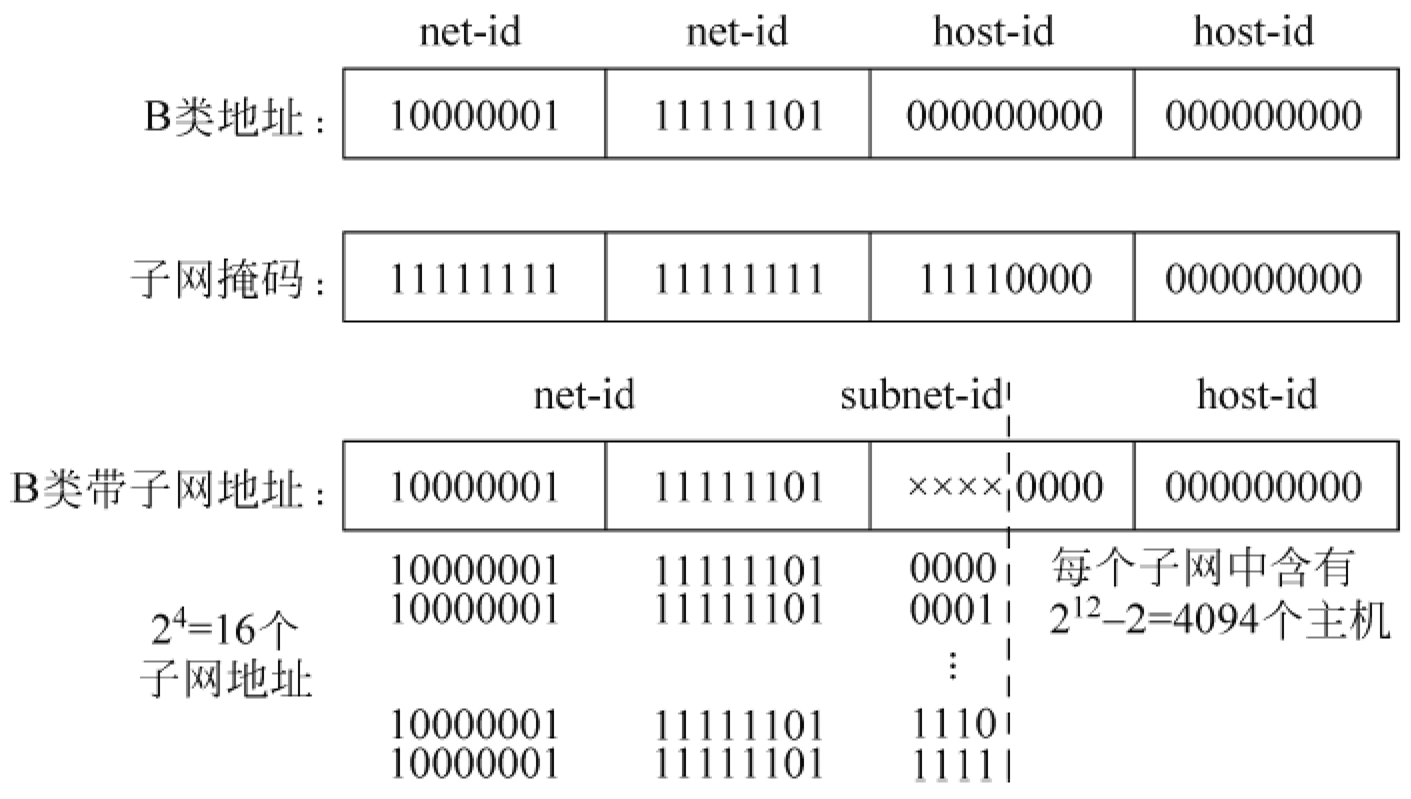


图 3.6　子网及主机分配

255.255.255.0,试给每一个地点分配一个子网号码,并计算出每个子网段地址的主机地址范围。

答:查表 3.2 可知,子网掩码为 255.255.255.0 的 B 类地址,当子网个数取 8 位时,可以有 254 个子网,如包含全 0 和全 1 子网时为 256 个子网,每个子网中含有 254 个主机,16 个子网就可以满足 4000(16×254=4096)台主机的要求。不足之处是还有 256−16=240 个子网处于闲置状态。以下给出的是具体分配情况:

0 号子网:129.253.0.0/24;主机地址:129.253.0.1～129.253.0.254

1 号子网:129.253.1.0/24;主机地址:129.253.1.1～129.253.1.254

2 号子网:129.253.2.0/24;主机地址:129.253.2.1～129.253.2.254

3 号子网:129.253.3.0/24;主机地址:129.253.3.1～129.253.3.254

4 号子网:129.253.4.0/24;主机地址:129.253.4.1～129.253.4.254

5 号子网:129.260.5.0/24;主机地址:129.253.5.1～129.253.5.254

6 号子网:129.260.6.0/24;主机地址:129.253.6.1～129.253.6.254

7 号子网:129.260.7.0/24;主机地址:129.253.7.1～129.253.7.254

8 号子网:129.260.8.0/24;主机地址:129.253.8.1～129.253.8.254

9 号子网:129.253.9.0/24;主机地址:129.253.9.1～129.253.9.254

10 号子网:129.253.10.0/24;主机地址:129.253.10.1～129.260.10.254

11 号子网:129.260.11.0/24;主机地址:129.260.11.1～129.260.11.254

12 号子网:129.260.12.0/24;主机地址:129.260.12.1～129.260.12.254

13 号子网:129.260.13.0/24;主机地址:129.260.13.1～129.260.13.254

14 号子网:129.260.14.0/24;主机地址:129.260.14.1～129.260.14.254

15 号子网:129.260.15.0/24;主机地址:129.260.15.1～129.260.15.254

【例 2】 某单位分配到一个地址块 136.23.12.96/27。现在需要进一步划分为 4 个一样大的子网。试回答以下 3 个问题。

(1) 每一个子网的网络前缀有多长?

答:由于 4 个子网需要 2 位,则每个网络前缀就是原掩码加 2,即 27+2=29 位。

(2) 每一个子网中有多少个地址?

答:由于主机位为 32−29=3,所以每个子网的地址中有 3 位留给主机用,因此共有

$2^3-2=6$ 个主机地址。

(3) 每一个子网的地址是什么？每一个子网可分配给主机使用的最小地址和最大地址是什么？

答：第一个地址块 136.23.12.96/29，可分配给主机使用的最小地址和最大地址如下。

最小地址：136.23.12.01100001＝136.23.12.97/29

最大地址：136.23.12.01100110＝136.23.12.102/29

第二个地址块 136.23.12.104/28，可分配给主机使用的最小地址和最大地址如下。

最小地址：136.23.12.01101001＝136.23.12.105/28

最大地址：136.23.12.01101110＝136.23.12.120/28

第三个地址块：136.23.12.112/28，可分配给主机使用的最小地址和最大地址如下。

最小地址：136.23.12.01110001＝136.23.12.113/28

最大地址：136.23.12.01110110＝136.23.12.119/28

第四个地址块：136.23.12.120/28，可分配给主机使用的最小地址和最大地址如下。

最小地址：136.23.12.01111001＝136.23.12.121/28

最大地址：136.23.12.01111110＝136.23.12.126/28

【例 3】 某公司有一个 C 类地址：222.222.5.0/24，要求用图示方法完成一个子网地址规划。

答：假设有 30 个部门，分成 $2^5-2=30$ 个子网，每个子网可配置 $2^3-2=6$ 个主机地址。这样，就需要把 IP 地址的最后一个八位组的主机地址分成两部分，其中子网部分占有 5 位，剩余 3 位为主机部分，则掩码为 24＋5＝29 位。（本题在子网规划中减去全 1 全 0 子网）

网络地址规划如图 3.7 所示，子网地址为 222.222.5.8/29～222.222.5.248/29，每个子网实际上可以用到 5 个主机地址，如 222.222.5.8 子网对应的主机地址为：222.222.5.10、222.222.5.11、222.222.5.12、222.222.5.13、222.222.5.14。另外，222.222.5.8 是主机部分全是 0 的 IP 地址，表示本子网的网络地址；222.222.5.15 是主机部分全是 1 的 IP 地址，是本子网的广播地址；222.222.5.9 是连接路由器的 IP 地址，也就是人们通常习惯上定义的“IP 网关地址”。

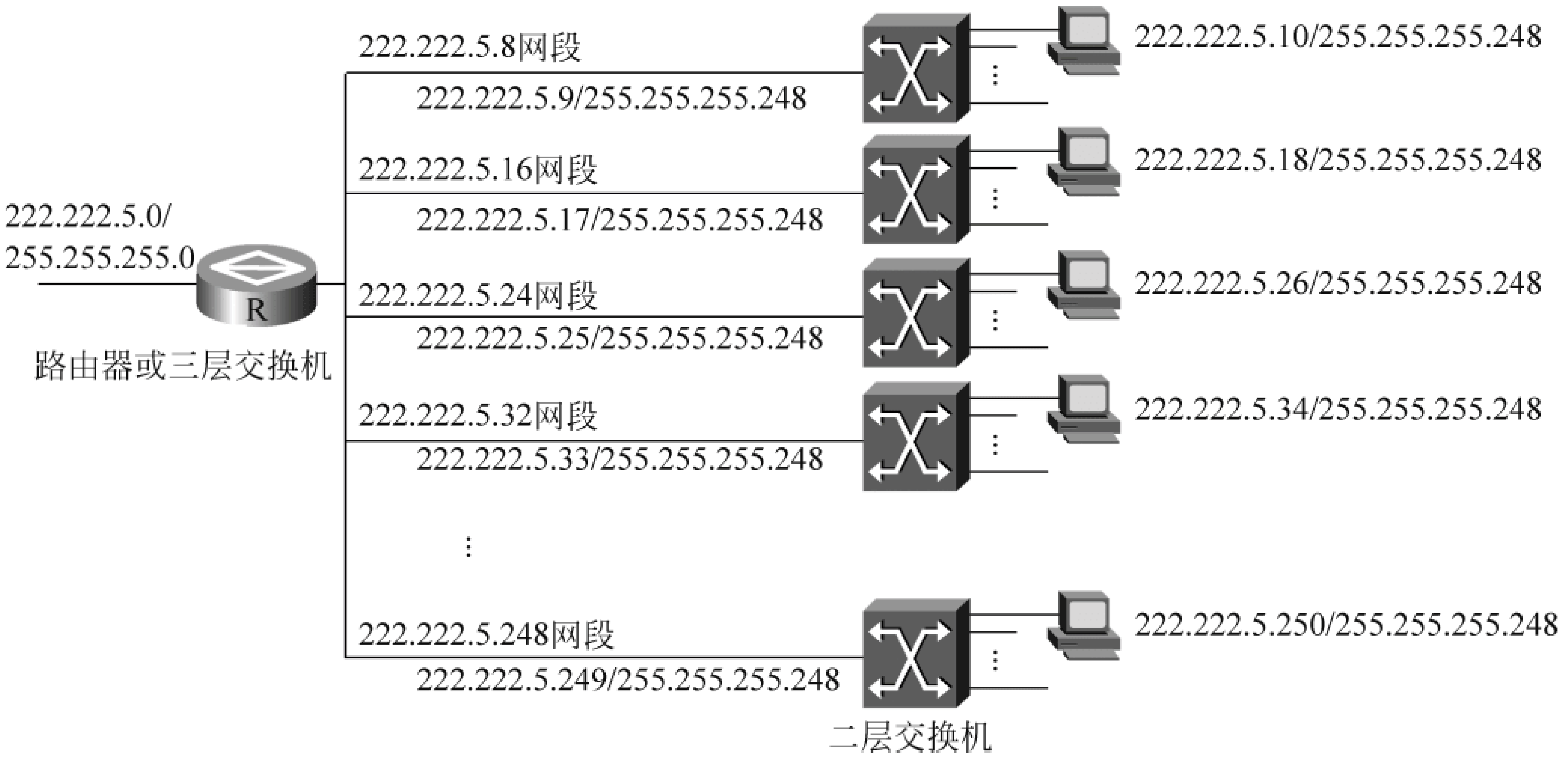


图 3.7　子网地址规划

对于 B 类网络来说，如果子网有 8 位，则能提供 254 个子网，每个子网可容纳 254 台主机。B 类地址在具体子网设计中可以查询表 3.2，表示的是位数与掩码等对应关系。对于 C 类子网地址来说，可查询表 3.3。

3.2.2 变长子网掩码

要求把一个网络划分成多个子网，但是每个子网的主机数不一定相同，而且相差很大，如果每个子网都采用固定长度子网掩码，而每个子网上分配的地址数相同，这就造成地址的大量浪费。这时，可以采用变长子网掩码(Variable Length Subnet Masking，VLSM)技术，对主机比较多的子网采用较短的子网掩码，子网掩码较短的地址可表示的网络/子网数较少掩码，而子网可分配的主机地址较多；主机数比较少的子网采用较长的子网掩码，可表示的网络/子网数较多掩码，而子网上可分配主机地址较少。这种寻址方案的好处是可节省地址。

【例 4】 如图 3.8 所示，要求对某公司用的 C 类网络 IP 地址 192.168.1.0 进行子网规划。这个公司共购置了 6 台路由器，1 台路由器作为企业网的网关路由器接入当地 ISP，其他 5 台路由器连接 5 个办公点，每个办公点 20 台 PC。

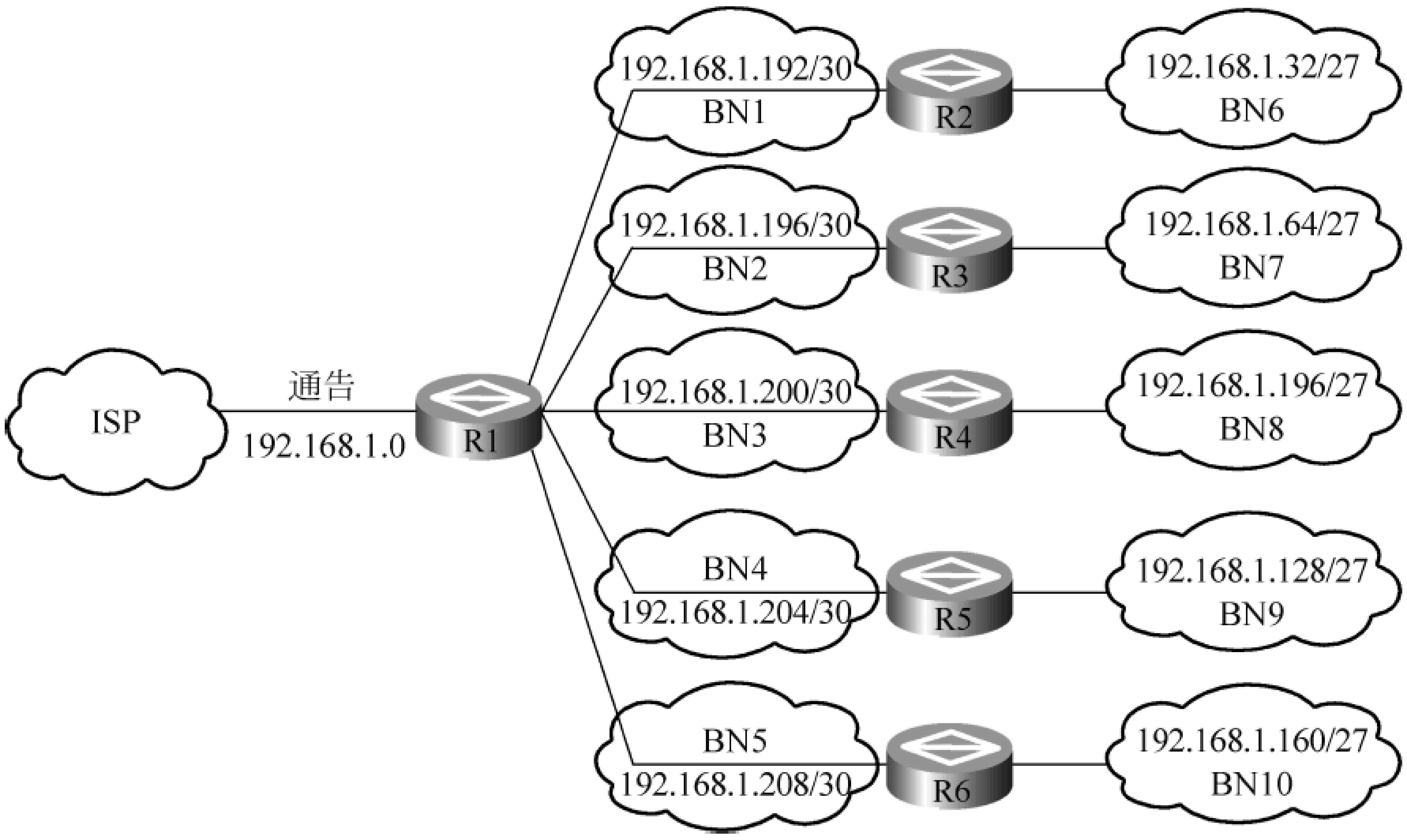


图 3.8　变长子网掩码示意

答：从图 3.8 中可以看出，需要划分 10 个子网：BN1～BN10，每个办公点网段需要 21 个 IP 地址(包括一个路由器接口)。办公点路由器与网关路由器相连的 5 个网段需要 2×5 个 IP 地址，每个网段 IP 地址数目差异较大，可以采用 VLSM 技术。5 个办公点网段采用子网掩码 255.255.255.224，划出 3 个子网位，办公点网段地址分别为 192.168.1.32、192.168.1.64、192.168.1.96、192.168.1.128、192.168.1.160，每个网段有 5 个主机位，可以容纳 $2^5-2=30$ 台主机，但实际上只能装 29 台主机，因为要去掉一个和路由器相连的网关地址。对于 5 个办公点路由器和网关路由器相连网段，划出 6 个子网位，网段地址分别为 192.168.1.192、192.168.1.196、192.168.1.200、192.168.1.204、192.168.1.208，子网掩码为 255.255.255.252，剩余 2 个主机位，每个子网最多有 2^2-2 个$=2$ 个合法 IP 地址，恰好每个路由接口分配一个 IP 地址。变长子网地址分配如图 3.9 所示，具体规划如下。

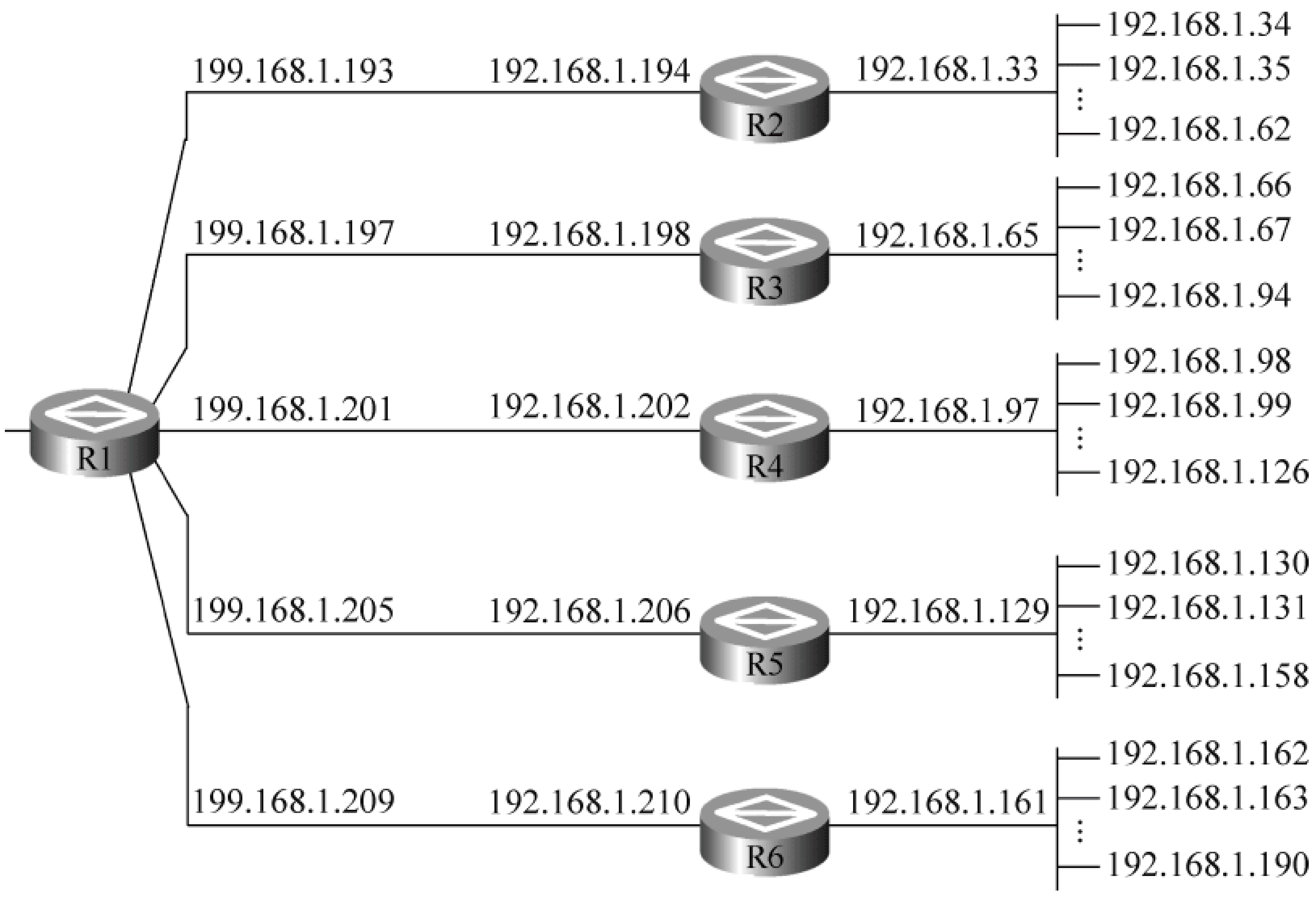


图 3.9　变长子网地址分配示意

(1) BN1 有效主机 IP 地址范围为 199.168.1.193～192.168.1.194。其中,R1 接口地址为 199.168.1.193;R2 接口地址为 192.168.1.194。

(2) BN2 有效主机 IP 地址范围为 199.168.1.197～192.168.1.198。其中,R1 接口地址为 199.168.1.197;R3 接口地址为 192.168.1.198。

(3) BN3 有效主机 IP 地址范围为 199.168.1.201～192.168.1.202。其中,R1 接口地址为 199.168.1.201;R4 接口地址为 192.168.1.202。

(4) BN4 有效主机 IP 地址范围为 199.168.1.205～192.168.1.206。其中,R1 接口地址为 199.168.1.205;R5 接口地址为 192.168.1.206。

(5) BN5 有效主机 IP 地址范围为 199.168.1.209～192.168.1.210。其中,R1 接口地址为 199.168.1.209;R6 接口地址为 192.168.1.210。

(6) BN6 有效主机 IP 地址范围为 192.168.1.33～192.168.1.62。其中,网关地址,即连接 R2 的接口地址为 192.168.1.33。

(7) BN7 有效主机 IP 地址范围为 199.168.1.65～192.168.1.94。其中,网关地址,即连接 R3 的接口地址为 192.168.1.65。

(8) BN8 有效主机 IP 地址范围为 199.168.1.129～192.168.1.158。其中,网关地址,即连接 R4 的接口地址为 192.168.1.97。

(9) BN9 有效主机 IP 地址范围为 199.168.1.97～192.168.1.126。其中,网关地址,即连接 R5 的接口地址为 192.168.1.129。

(10) BN10 有效主机 IP 地址范围为 199.168.1.161～192.168.1.190。其中,网关地址,即连接 R6 的接口地址为 192.168.1.161。

3.2.3　无类域间路由

无类域间路由(Classless Inter-Domain Routing,CIDR)由 RFC 1817 定义。CIDR 放弃了传统 IP 地址分类边界,把路由表中的若干条路由汇聚为一条路由,减少了路由表的规模,

提高了路由器的可扩展性。

CIDR 举例如图 3.10 所示,一个 ISP 被分配了一些 C 类网络:199.188.0.0~199.188.255.0。ISP 准备把这些 C 类网络分配给各个用户群,目前已经分配了 3 个 C 类网段给用户。如果没有实施 CIDR 技术,ISP 的路由器的路由表中会有 3 条下连网段的路由条目,并且会把它通告给 Internet 上的路由器。通过实施 CIDR 技术,就可以在 ISP 的路由器上把这 3 个网段 199.188.1.0/24、199.188.2.0/24、199.188.3.0/24 汇聚成一条路由 199.188.0.0/16。这样 ISP 路由器只向 Internet 通告 199.188.0.0/16 这一条路由,从而减少了路由表的数目。

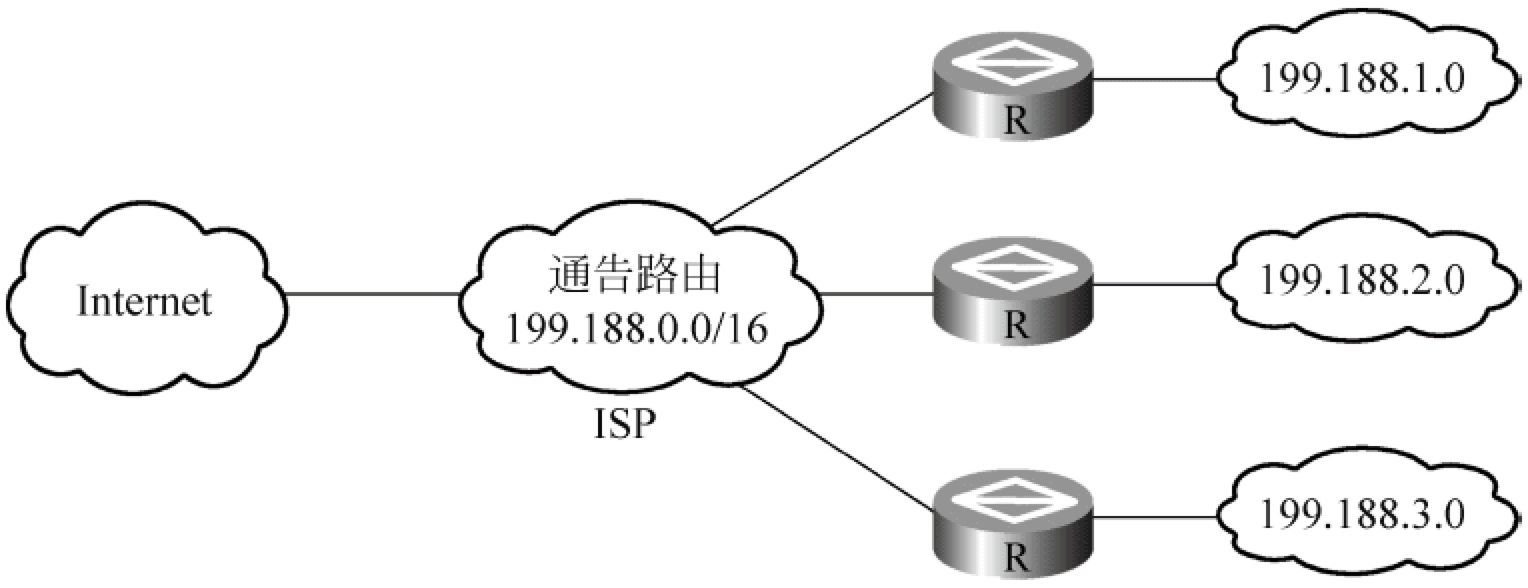


图 3.10 CIDR 举例

在使用 CIDR 技术汇聚的网络地址时,要求对应位必须是一致的,如果将图 3.10 所示的 ISP 连接了一个 177.186.1.0/32 网段,这个网段路由将无法汇聚、无法实现 CIDR 技术。

3.3 Internet 应用协议

Internet 是一个通用名词,泛指由多个计算机网络互联而成的虚拟网络。Internet 又分为企业内部网(intranet)、外部网(extranet)、万维网(WWW 或 Web)等。Internet 的应用层提供了多种标准化的应用协议,是直接由 TCP/UDP/SCTP 支持的。

3.3.1 DNS(域名服务器)

域名服务器(DNS),专门从事域名和 IP 地址之间的转换翻译工作。实际上域名服务器相当于一本电话簿,已知姓名就可以查到电话号码。域名地址本身是分级结构的,所以域名服务器也是分级的。域名服务器相当于一个数据库,它存储着一定范围内的主机和网络的域名及相应的 IP 地址。

域名系统的主要功能:将域名解析为主机能识别的 IP 地址。Internet 上的域名服务器系统也是按照域名的层次来安排的。每个域名服务器都只对域名体系中的一部分进行管辖。共有 3 种不同类型的域名服务器,即本地域名服务器、根域名服务器、授权域名服务器。

使用 IP 地址很不方便,因此 Internet 采用了一种字符型命名方法,即用表示一定意思的字符串来标识主机地址,两者相互对应,当然主机名也要保持全网统一。

域名通常就是指一个管理、维护一组主机和名字的系统,域名系统如图 3.11 所示。域名系统是一个分层的名字管理、查询系统,主要提供 Internet 上主机 IP 地址和主机名相互对应关系的服务。域名系统由网络信息中心(NIC)将主机名字空间划分为若干部分,并将

各部分的管理权授予相应的机构，各管理机构可以将管辖内的名字空间进一步划分成若干子部分，并将子部分的管理权再授予相应的子机构，以完成所属主机名和主机 IP 地址的管理。

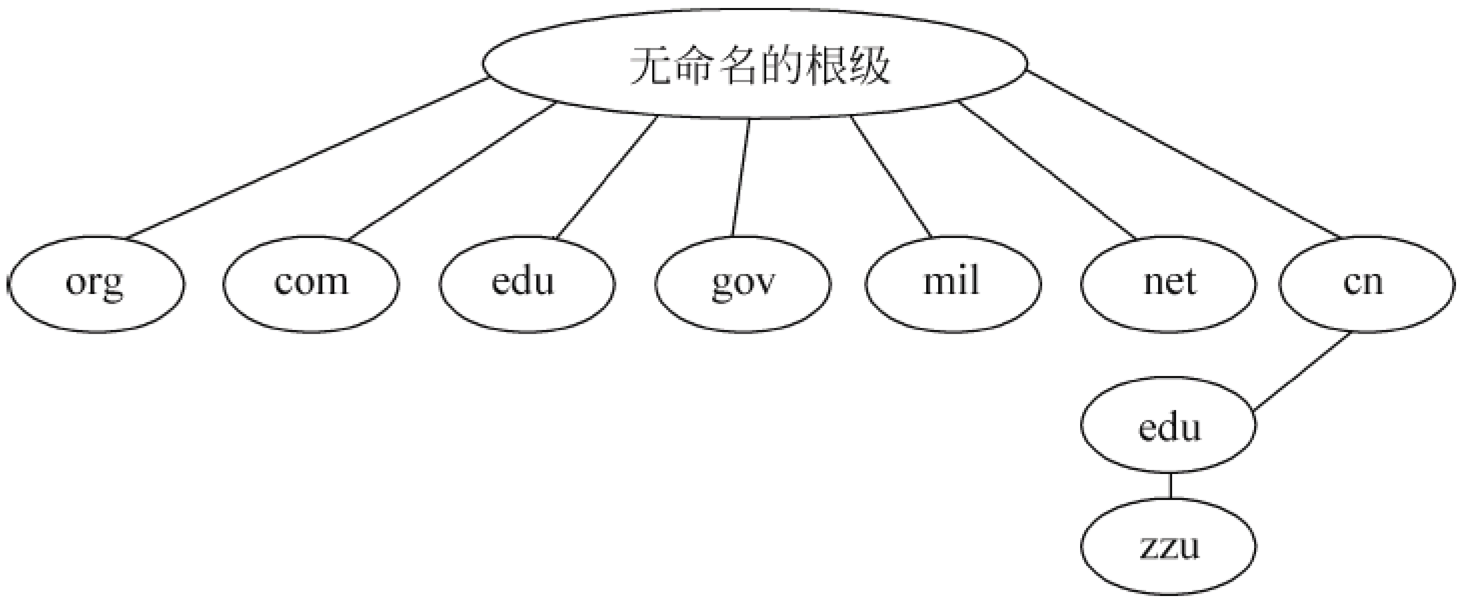


图 3.11 域名系统

Internet 的域名系统主要由多级域名组成：无命名的根级、一级、二级及各子级域名。Internet 的名字空间为树状结构，树上的每一节点为一个确定指示的域。

(1) 根级：为一特殊、由 NIC 管理的域，未命名。

(2) 一级域名：由根级(即 NIC)授权管理。一级域名通常有两种命名方式，一是按行业命名；二是按国家和地区命名，详见表 3.4。其中 3 个字符表示的域名对应于各部门，2 个字符是按国别、地理位置划分的国家域名，如 cn 表示中国。

表 3.4 一级域名的命名及含义

域名	com	edu	gov	mil	net	org
含义	商业组织	教育组织	政府部门	军事部门	主要网络	各种组织

(3) 二级域名：是一级域名的进一步划分，如 cn 下又可参照一级域名中的行业分类再分为 com、edu 等。

(4) 子级域名：子级域名是二级域名的进一步划分，子级域名可小到只管理一台主机，也可大到包含许多主机和进一步授权管理的子域，如郑州大学域名 zzu 就是中国 cn 的教育机构 edu 下的一个子域。

一个完整的域名，就是从根级域到当前域的所有组织名从右到左由“.”分隔符连接。如郑州大学的完整域名为 zzu.edu.cn，其中：cn 为一级域名，edu 为 cn 下的二级域名，zzu 为 edu.cn 的子级域名。当郑州大学申请到 zzu.edu.cn 域名后，就可对郑州大学内所有主机的主机名进行管理，而完整的主机名应为：主机名.域名。例如，校内有两台叫作 fred 和 cree 的主机，分属于化学系(chem)和物理系(phy)，其完整的主机名分别为：

```
frde.chem.zzu.edu.cn
cree.phy.zzu.edu.cn
```

机器名、域名全网必须唯一，但不同域名下可以使用相同的名字，如上例中两台机器都命名为 fred，但分属化学系和物理系，这是可以的。

```
fred.chem.zzu.edu.cn
fred.phy.zzu.edu.cn
```

当一个当地域名服务器不能立即回答某个主机的查询时，该当地域名服务器就以 DNS 客户的身份向某一个根域名服务器查询。若根域名服务器有被查询主机的信息，就发送 DNS 回答报文给本地域名服务器，然后当地域名服务器再回答发起查询的主机。

通常根域名服务器用来管辖顶级域，当根域名服务器没有被查询的主机的信息时，它一定知道某个保存有被查询的主机名字映射的授权域名服务器的 IP 地址。根域名服务器虽然不直接对顶级域下面所属的所有的域名进行转换，但它能够逐级找到下面的所有二级域名的域名服务器。每一个主机都会在授权域名服务器处注册登记，一个主机的授权域名服务器就是它的主机 ISP 的一个域名服务器。授权域名服务器总是能够将其管辖的主机名转换为该主机的 IP 地址。Internet 允许各自根据本单位的情况将本域名划分为若干个域名服务器管辖区。

【例 5】 说明域名到 IP 地址的解析过程。

答：当某一个应用进程需要把主机名解析为 IP 地址时，该应用进程就调用应用程序，并作为 DNS 的一个用户，把待解析的域名放在 DNS 请求报文中，以 UDP 用户数据报方式发给当地域名服务器，当地域名服务器再逐级查找域名后，把对应的 IP 地址放在回答报文中返回。应用程序获得目的的 IP 主机后即可进行通信。

【例 6】 说明 Internet 中通过域名的寻址过程，一个国外用户寻找一台叫 host.edu.cn 的中国主机，其过程如图 3.12 所示。

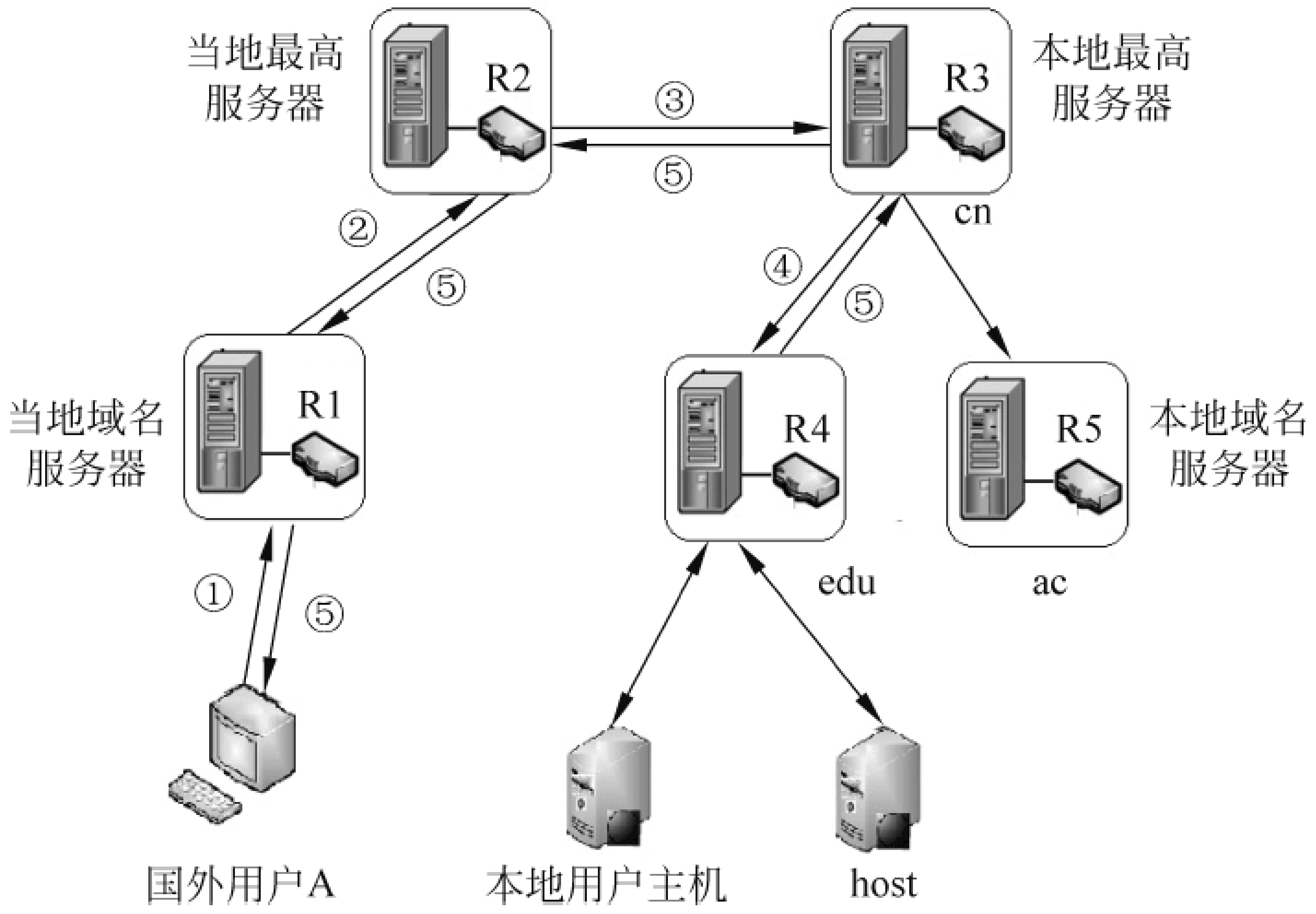


图 3.12　通过域名的寻址过程示意图

答：

(1) 国外用户 A“呼叫”host.edu.cn，当地域名服务器受理并分析号码(经路由器 R1)。

(2) 由于当地域名服务器中没有中国域名资料，必须向上一级查询，图中当地域名服务器向当地最高域名服务器问询(经路由器 R2)。

(3) 当地最高域名服务器检索自己的数据库，查到 cn 为中国的一级域名，则指向中国

的本地最高域名服务器(经路由器 R3)。

(4) 中国最高域名服务器分析号码,看到第 2 级域名为 edu,就指向本地的 edu 域名服务器,从图中可以看到 ac 域名服务器与 edu 域名服务器是平级的(经路由器 R5、R4)。

(5) 经过 edu 域名服务器分析,看到第 3 级域名是 host,就将名为 host 主机的 IP 地址返送给 A。

3.3.2 FTP(文件传输协议)

文件传输协议(File Transfer Protocol,FTP)包括两个组成部分,其一为 FTP 服务器,其二为 FTP 客户机。其中 FTP 服务器用来存储文件,用户可以使用 FTP 客户机通过 FTP 协议访问位于 FTP 服务器上的资源。在开发网站的时候,通常利用 FTP 协议把网页或程序传到 Web 服务器上。默认情况下,FTP 协议使用 TCP 端口 20 传输数据,使用 21 端口传输控制信息。但是,是否使用 20 作为传输数据的端口与 FTP 使用的传输模式有关,如果采用主动模式,那么数据传输端口就是 20;如果采用被动模式,则具体最终使用哪个端口还需要由服务器端和客户机协商决定。图 3.13 给出了使用 FTP 的文件传送过程,具体工作流程如下:

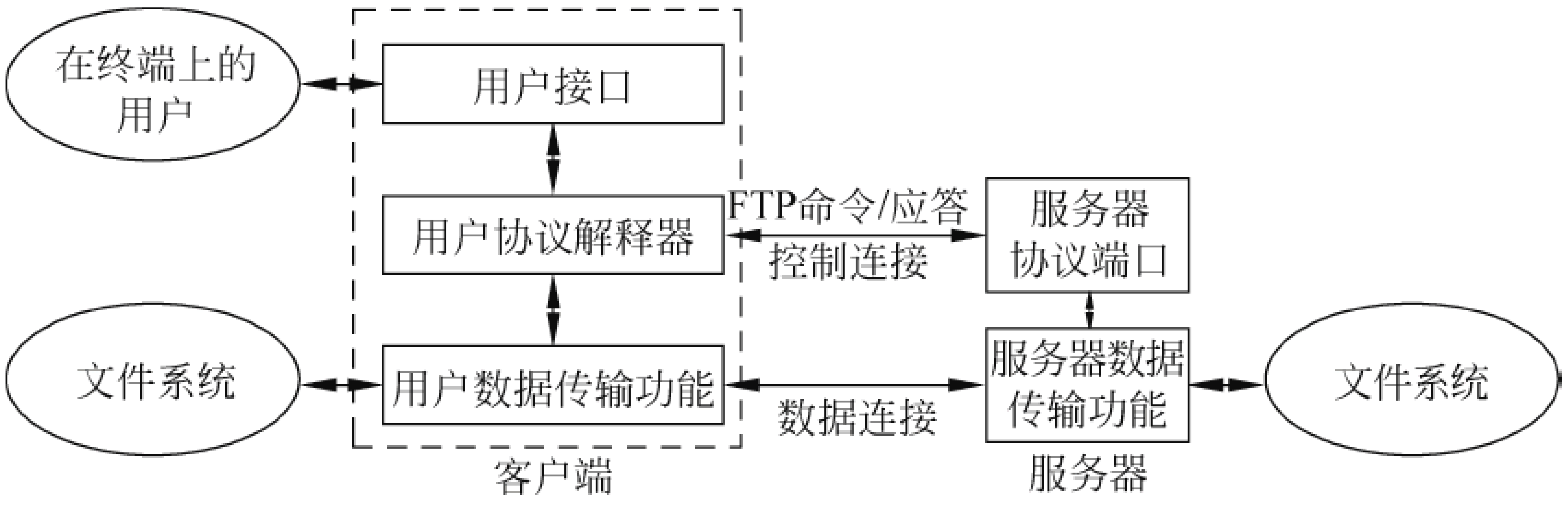


图 3.13 文件传输中的处理过程

(1) 服务器协议接口打开熟知端口(端口号为 21),使客户端发来的进程能够顺利连接。

(2) 等待用户协议解释器客户进程发出 FTP 连接请求命令。

(3) 当服务器收到客户进程请求时,启动从属进程来处理,并回送 FTP 应答,从属进程对客户进程处理完毕之后即终止。在客户发来的连接请求里还提供自己用来传送数据的另一个端口号。

(4) 服务器(数据传送功能)用自己的熟知端口(端口号为 20)与客户(数据传送功能)进程提供的端口号建立数据传送连接,并开始 FTP 数据传送。

(5) 服务器协议端口回到等待状态,继续接收其他客户进程发来的请求。主进程与从属进程的处理是可以并发产生的。

【例 7】 为什么说 FTP 是带外传送控制信息?主进程和从属进程各起什么作用?

答:

(1) 在进行文件传输时,FTP 的客户与服务器之间要建立两个并行的 TCP 连接:“控制连接”和“数据连接”。控制连接不发送文件,而是控制整个进程,实际传输文件的是数据连接。因此,FTP 使用的是一个分离的控制连接,因此 FTP 的控制信息是带外传输的。

(2) FTP 的服务进程由两大部分组成:一个主进程负责接收新的请求;另外有若干从属进程,负责处理单个请求。

3.3.3 Telnet(远程登录协议)

Telnet 是一个简单的远程终端协议,是一个客户机/服务器模式的协议,用户用 Telnet 就可以将其当地主机通过 TCP 连接注册到远地的另一个主机上,Telnet 能将当地用户的击键传到远地主机,同时也能将远地主机的输出通过 TCP 连接返回到当地用户屏幕。Telnet 协议的目的是提供一个相对通用的、双向的通信方法,能通过一个标准过程进行互相交互。

【例 8】 结合图 3.14 说明客户机/服务器模式的 Telnet 协议,Telnet 协议的主要特点是什么? 为什么需要网络虚拟终端(Network Virtual Terminal,NVT)?

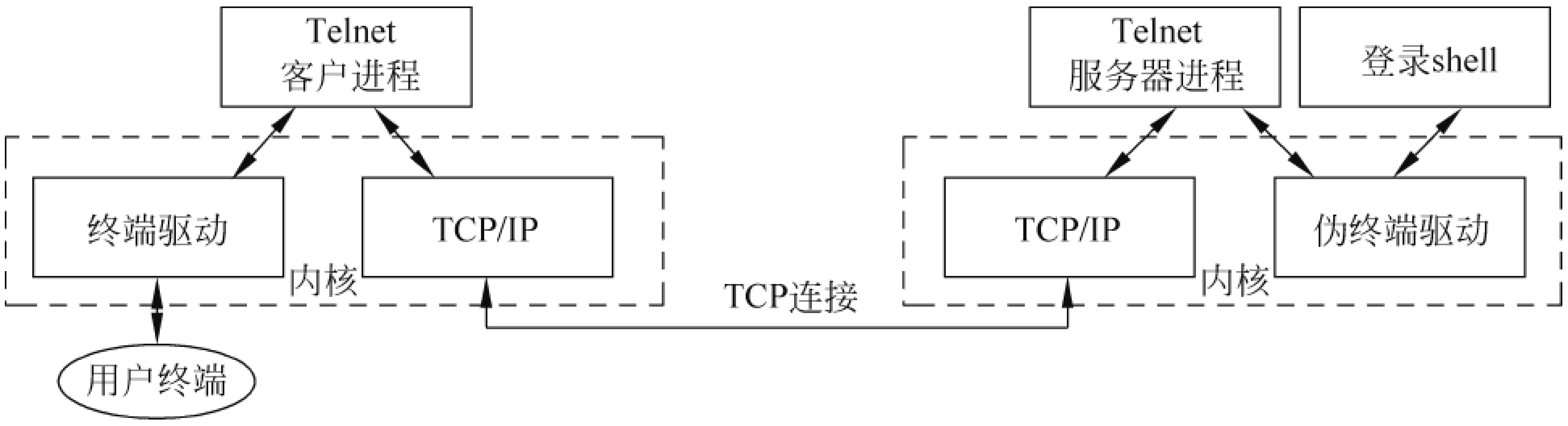


图 3.14 客户机/服务器模式的 Telnet 简图

答:Telnet 主要特点如下:

(1) 当地用户终端运行 Telnet 客户机程序,远程主机运行 Telnet 服务器程序。客户机与服务器之间执行 Telnet NVT 协议,在两端执行各自的操作系统功能。

(2) Telnet 服务器可以应付多个并发的连接,客户进程与服务器的固定端口(23)建立 TCP 连接,实现 Telnet 服务。

(3) Telnet 协议是 TCP/IP 协议族中的一员,是 Internet 远程登录服务的标准协议。应用 Telnet 协议能够把当地用户所使用的计算机变成远程主机系统的一个终端。

(4) Telnet 定义一个网络虚拟终端为远程系统提供一个标准接口。客户机程序不必详细了解远程系统,它们只需构造使用标准接口的程序。

(5) Telnet 包括一个允许客户机和服务器协商选项的机制,而且它还提供一组标准选项。

(6) Telnet 对称处理连接的两端,即 Telnet 不强迫客户机从键盘输入,也不强迫客户机在屏幕上显示输出。

为了适应异构环境,Telnet 定义了数据和命令应怎样通过 Internet,这些定义就是所谓的网络虚拟终端 NVT。客户软件把用户的击键和命令换成 NVT,并交送给服务器。Telnet 协议定义了数据和命令在 Internet 上的传输方式,它的应用过程如下:

(1) 对于发送的数据,客户机软件把来自用户终端的按键和命令序列转换为 NVT 格式,并发送到服务器,服务器软件将收到的数据和命令,从 NVT 格式转换为远地系统需要的格式。

(2) 对于返回的数据,远地服务器将数据从远地机器的格式转换为 NVT 格式,而本地客户机将接收到的 NVT 格式数据再转换为本地的格式。

3.3.4 SMTP(简单邮件传送协议)

图 3.15 给出了电子邮件主要组成构件,主要用到邮局协议(Post Office Protocol,POP)和简单邮件传输协议(Simple Mail Transfer Protocol,SMTP)两个协议。其中:

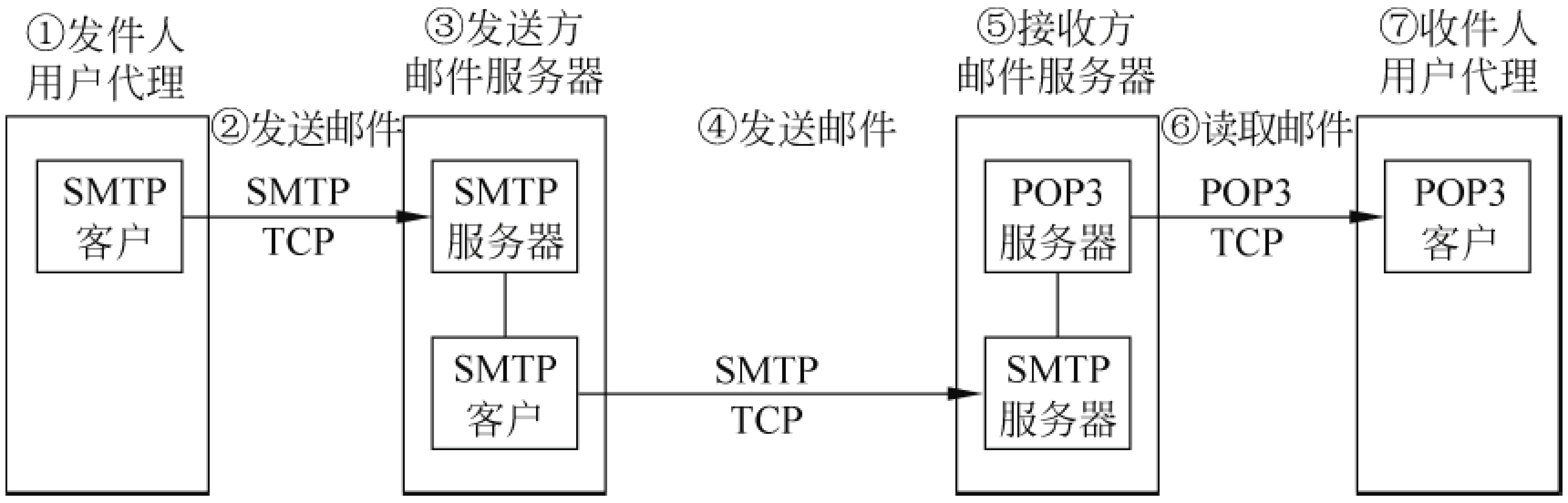


图 3.15 电子邮件主要组成构件

SMTP 是一组用于由源地址到目的地址传送邮件的规则,由它来控制信件的中转方式。SMTP 则是负责邮件服务器之间的通信协定,为每台计算机在发送或中转信件找到下一个目的地。

POP 是用于电子邮件的接收,主要支持使用客户端远程管理在服务器上的电子邮件。POP 服务器是用来收信的,而且每个 E-mail 地址一般只有一个。如果需要同时收取多个邮箱的信件,就必须要设置每个邮箱的 POP3 服务器地址。

在接收邮件的用户 PC 机中必须运行 POP 客户机程序,而在其 ISP 的邮件服务器中则运行 POP 服务器程序。POP 服务器只有在用户输入鉴别信息(用户名和口令)后才允许对邮箱进行读取。POP 是一个脱机协议,所有对邮件的处理都在用户的 PC 机上进行。

【例 9】 结合图 3.15 给出的电子邮件主要组成构件,简述 POP 和 SMTP 协议,邮局协议 POP3 指的会话过程为哪 3 个状态?SMTP 通信有哪几个阶段?说明电子邮件的格式各部分的意思。

答:

(1) POP3 是指它的会话过程分为 3 个状态:验证状态、事务状态和更新状态。

① 验证状态是在建立 TCP 连接后,客户端将认证信息(用户名,密码)传送给服务器。

② 事务状态是通过发送相关的命令让服务器进行邮件事务处理。

③ 更新状态是由服务器释放所有资源,结束与客户端的连接。

(2) SMTP 通信过程如图 3.15 所示,按图中标号从①到⑦流程,主要概括为以下 3 个阶段:

建立 TCP 连接,连接是在发送主机的 SMTP 客户和接收主机的 SMTP 服务器之间建立的,不使用中间的邮件服务器;传送邮件阶段;邮件发送完毕后,SMTP 应释放 TCP 连接。

(3) TCP/IP 体系的电子邮件系统规定电子邮件地址的格式如下:

收信人邮箱名@邮箱所在主机的域名,其中符号"@"表示"在"的意思。

3.3.5 WWW(万维网)

万维网(World Wide Web,WWW),又称 3W 或 Web,它是一种基于超文本和 HTTP 的、全球性的、动态交互的、跨平台的分布式图形信息系统。WWW 的信息可以是文字、图

形图像、声音、动画等类型，用户可以使用 WWW 客户程序方便地浏览、检索和查找所需要的信息。WWW 具有如下特点：

(1) 采用超文本链接的信息系统：用户在浏览 WWW 过程中，可以通过超文本链接从一处跳转到另一处。

(2) 拥有图形界面：WWW 既能展示文字，又能提供集图像、声音、视频和文字信息。

(3) 不受操作平台和硬件设备的限制：只要系统进入 WWW，不管拥有何种计算机硬件设备，安装的是何种操作系统，使用 Web 浏览器，便可以获取所需的信息。

(4) 采用分布式信息系统。大量的信息存储分布在全球的各个 Web 网站上，用户可以从这些站点上获取所需要的信息。

(5) 采用动态网页：由于 Web 页存储在发布它的站点上的，所以网站管理人员可以动态地、及时地更新信息传播内容及方式。

(6) 采用交互式的浏览方式：WWW 网上的信息传递可以是双向进行。

【例 10】 要求对万维网常用到的一些名词进行解释。

(1) 统一资源定位符(Uniform Resource Locator，URL)：是用来表示从 Internet 上得到的资源位置和访问这些资源的方法。

(2) 超文本传送协议(HyperText Transfer Protocol，HTTP)：使万维网客户程序与万维网服务器程序之间的交互遵守严格的协议。

(3) 超文本标记语言(HyperText Markup Language，HTML)：使得 WWW 页面的设计者可以很方便地用链接从本页面的某处链接到 Internet 上的任何一个 WWW 页面，并且能够在自己的主机屏幕上将这些页面显示出来。

(4) 通用网关接口(Common Gateway Interface，CGI)：是一种标准，它定义了动态文档应如何创建，输入数据应如何提供给应用程序，以及输出结果应如何使用。

(5) 浏览器：是指可以显示网页服务器或者文件系统的 HTML 文件(标准通用标记语言的一个应用)内容，并让用户与这些文件交互的一种软件。一个浏览器包括一组客户程序、一组解释程序，以及一个控制程序。

(6) 超文本：包含指向其他文档的链接的文本。

(7) 超媒体：除包含指向其他文档的链接的文本，还包含用其他方式表示的信息，如：图像、图形、声音、动画，甚至活动视频图像。

(8) 超链：也叫链接，在文档中有一些地方的文字是用特殊方式显示的，当我们将鼠标移动到这些地方时，鼠标箭头就变成一只手的形状，如果我们在这些地方点击鼠标，就可以从这个文档链接到另一个文档。

(9) 页面：在一个客户程序主窗口上显示出的 WWW 文档。

(10) 活动文档：一种提供屏幕连续更新的技术，这种技术是把所有的工作都转移给浏览器端。每当浏览器请求一个活动文档时，服务器就返回一段活动文档程序副本，使该程序副本在浏览器端运行。只要用户运行活动文档程序，活动文档的内容就可以连续地改变。由于活动文档技术不需要服务器的连续更新传送，对网络带宽的要求也不会太高。

(11) 搜索引擎(search engine)：在 WWW 中用来进行搜索的工具叫作搜索引擎。大体可分为全文检索搜索引擎和分类目录搜索引擎。

(12) 网络爬虫：是一个自动提取网页的程序，它为搜索引擎从 WWW 上下载网页，是

搜索引擎的重要组成。按照系统结构和实现技术可以分为以下几种类型：通用网络爬虫(general purpose web crawler)、聚焦网络爬虫(focused web crawler)、增量式网络爬虫(incremental web crawler)、深层网络爬虫(deep web crawler)。

【例 11】 假设某主机从已知的 URL 获得一个 WWW 文档，而开始时并不知道该 WWW 服务器的 IP 地址。试问：除 HTTP 外，还需要什么应用层协议和传输层协议？

答：不知道 WWW 服务器的 IP 地址，在应用层协议应该调用 DNS 获取。

传输层协议需要的是 UDP(支持 DNS 使用)和 TCP(支持 HTTP)。

【例 12】 基于 WWW 的电子邮件系统有什么特点？在传送邮件时是用什么协议？

答：特点：无论何时何地，只要能上网，在打开 WWW 浏览器之后，就可以收发电子邮件。这时，邮件系统中的用户代理就是普通的 WWW。

用的协议：HTTP 和 SMTP。比如，电子邮件从主机 A 发送到新浪邮件服务器(发信信箱)是使用的 HTTP 协议；两个邮件服务器之间的传送是使用 SMTP；邮件从网易邮件服务器(收信信箱)传送到主机 B 是使用 HTTP 协议。

3.3.6 DHCP(动态主机配置协议)

动态主机配置协议(Dynamic Host Configuration Protocol，DHCP)是一个局域网的网络协议，传输层使用 UDP 协议，主要用途是给内部网络或网络服务供应商自动分配 IP 地址。DHCP 有 3 个端口，其中 UDP67 为 DHCP Server(DHCP 服务器)服务端口；UDP68 为 DHCP Client(DHCP 客户机)的服务端口，而 UDP546 则用于 DHCPv6 Client。

【例 13】 结合图 3.16 说明为何引入 DHCP Relay 的工作原理。

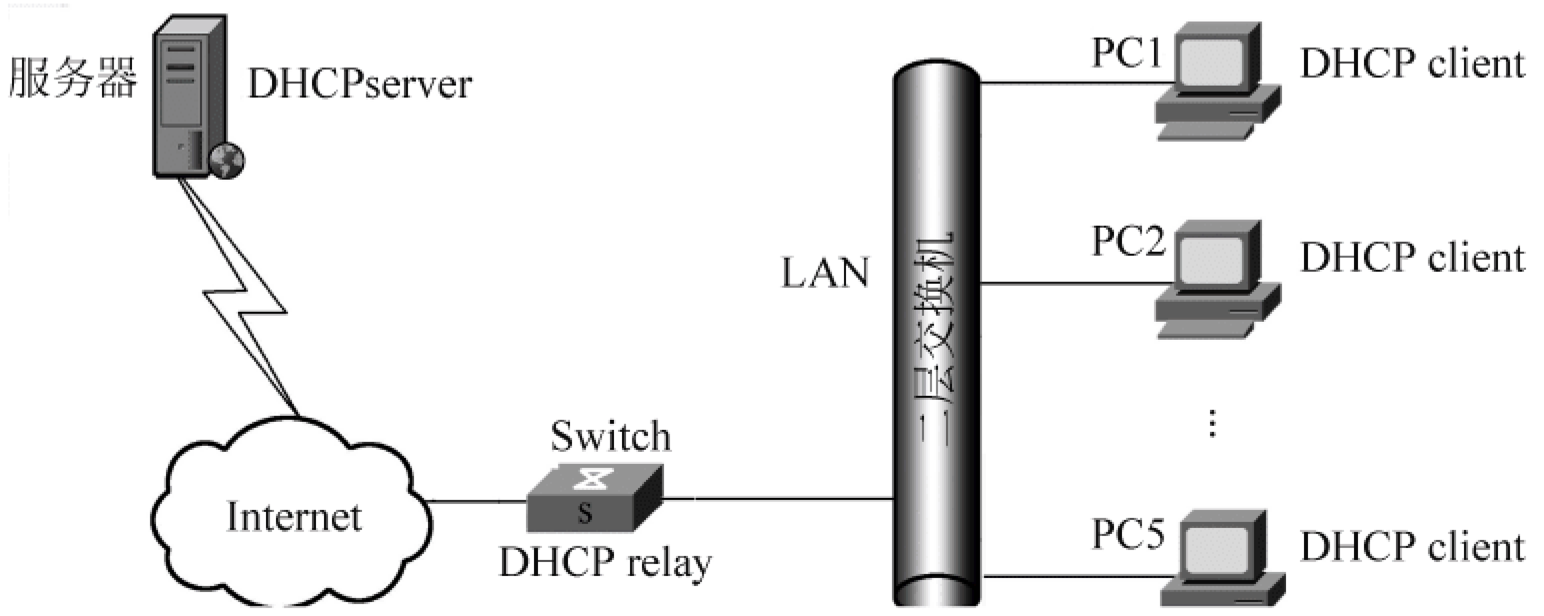


图 3.16 远端设置 DHCP Server 网络架构

答：引入了 DHCP relay(DHCP 中继，也称 DHCP 中继代理)：DHCP relay 可以实现在不同子网和物理网段之间处理和转发 DHCP 信息的功能。局域网内的 DHCP 客户端可以通过 DHCP relay 与其他子网的 DHCP 服务器通信，最终取得合法的 IP 地址。这样，多个网络上的 DHCP 客户端可以使用同一个 DHCP 服务器，既节省了成本，又便于进行集中管理。DHCP relay 的典型应用示意图如图 3.16 所示。

DHCP relay 工作原理如下：

(1) 当 DHCP client 启动并进行配置初始化时，它会在本地网络广播配置请求报文。

(2) 如果本地网络存在 DHCP server，则可以直接进行 DHCP 配置，不需要 DHCP relay。

(3) 如果本地网络没有 DHCP server，则与本地网络相连的具有 DHCP relay 功能的网络设备收到该广播报文后，将其转发给指定的其他网络上的 DHCP server。

(4) DHCP server 根据 DHCP client 提供的信息进行相应的配置,并通过 DHCP relay 将配置信息发送给 DHCP client,完成其动态配置。

【例 14】 在什么情况下使用 DHCP 协议?当一台计算机第一次运行引导程序时,其 ROM 中有没有该主机的 IP 地址、子网掩码或某个域名服务器的 IP 地址?

答: 动态主机配置协议 DHCP 提供了即插即用连网的机制,这种机制允许一台新加入网络的计算机获取 IP 地址,而不需要人为参与。在无盘计算机的情况下,操作系统和连网软件可以存储在 ROM 中。但是制造厂家并不知道主机的 IP 地址、子网掩码等信息,这些信息取决于该机器所连接到的网络,因此在 ROM 中没有存储这些信息。

【例 15】 何为动态文档?

答: 动态文档(dynamic document)是与 WWW 文档有关的计算程序。当浏览器需要动态文档时,服务器就运行该程序并发送输出到浏览器。动态文档程序对每个不同的需求可生成不同的输出。

3.3.7 SNMP(简单网络管理协议)

简单网络管理协议(Simple Network Management Protocol, SNMP)是用来管理 Internet 上众多厂家生产的软硬件平台,以及接入 Internet 的其他通信设备,它的基本功能包括监视网络性能,检测分析网络差错和网络配置管理等,并将检测到的问题发送到网络管理工作站。

图 3.17 为典型的 SNMP 管理模型,整个系统必须要有一个管理站(management station),它实际上是网控中心,是由网管人员在维护操作,在每个被管对象中一定要有代理进程(proxy agent)设备,就是将被管理的设备信息都被放置到这里,配置有信息管理库管理信息库 MIB,只有在 MIB 中的对象才是 SNMP 可以管理的;管理进程和代理进程利用 SNMP 报文进行通信,SNMP 管理程序使用 UDP(162),SNMP 代理程序使用 UDP(161)来传送 SNMP 报文。

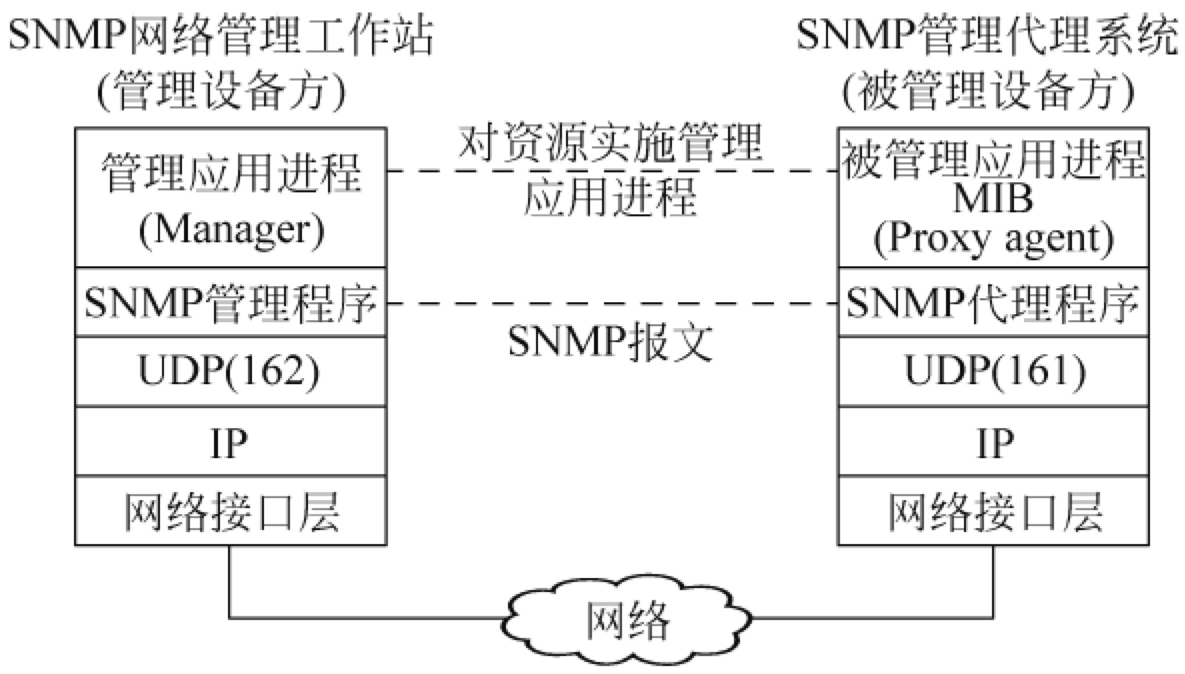


图 3.17 SNMP 管理模型

【例 16】 什么是管理信息库 MIB?它的作用是什么?

答: 管理信息库(Management Information Base, MIB)是一个网络中所有可能的被管对象集合的数据结构。MIB 的定义与具体的网络管理协议无关,这对于厂商和用户都有利,厂商可以在产品中包含 SNMP 代理软件,并保证在定义新的 MIB 项目后该软件仍能够遵守标准。用户可以使用同一网络管理客户软件来管理具有不同版本的 MIB 的多个路由器。

MIB 的作用如下。

(1) MIB 用来描述树的层次结构，它是所监控网络设备的标准变量定义的集合。

(3) MIB 定义了命名对象和定义对象类型(包括范围和长度)的通用规则。

(3) MIB 定义了存储被管对象的数据类型的种类。

(4) MIB 定义了管理对象及其值进行编码的规则，解决传送的管理数据如何编码的问题。

3.4 Internet 接入技术

3.4.1 CHINANET 与接入技术概述

1. CHINANET

CHINANET 是中国公用计算机互联网，采用 TCP/IP 协议，并通过高速数字专线与国际 Internet 互连。CHINANET 与国内的企业网、校园网和各种局域网互连，构成中国的 Internet。CHINANET 与其他公用数据网和公用电话网互连，可以向所有客户提供 Internet 服务。

CHINANET 由骨干网和接入网组成，并设立全国网管中心和接入网网管中心。骨干网是 CHINANET 主要信息通路，主要负责转接全网的业务，并为接入网提供接入端口。CHINANET 最初设置两条国际电路实现与国际 Internet 的互连，国际出入口节点分别设立在北京和上海。

由于 CHINANET 将电话网、DDN、分组交换网、帧中继网等多种公用通信网互连在一起，所以 CHINANET 能为用户提供多种连接方式，如图 3.18 所示。

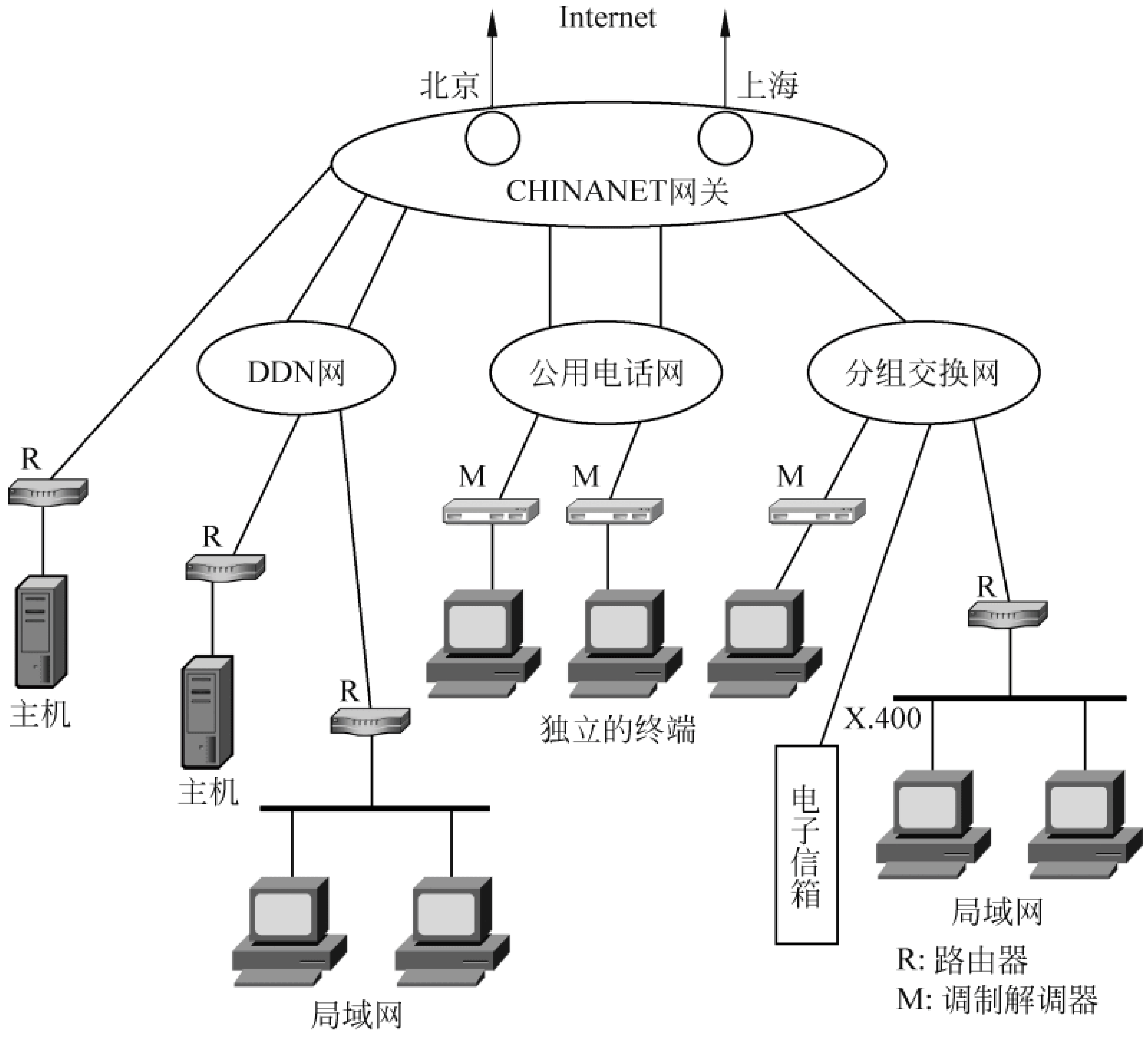


图 3.18 CHINANET 接入示意图

(1) 终端或局域网通过模拟或数字专线,采用 TCP/IP 接入 Internet。

(2) CHINADDN 网的终端或局域网用户利用 CHINADDN 或帧中继协议,通过 CHINADDN 网和相应的网关接入 Internet。

(3) 分组交换网的终端、局域网用户专线或电话拨号方式,以 X.25 协议通过分组交换网和相应的网关接入 Internet。

(4) 个人终端用户通过公用电话网,采用拨号方式,以终端仿真方式,或利用串行线路网间协议及点到点协议,通过终端服务器接入 Internet。

CHINANET 最大的好处是将分组交换网、电话网与 DDN 的用户互连在一起,直接实现了在 Internet 上的"对接",CHINANET 为公众提供了各种接入方式。

2. 接入技术概述

接入网可分为有线接入网和无线接入网(见表 3.5),有线接入网包括铜线接入网、光纤接入网和混合光纤/同轴电缆接入网。目前,在有线接入方面,出现了以双绞线为基础的高比特率数字用户环路系统(HDSL)、非对称数字环路系统(ADSL)等,以光纤为基础的光纤到小区(FTTZ),光纤到路边(FTTC),光纤到大楼(FTTB),最终实现光纤到户(FTTH)的接入系统。还有混合光纤/同轴电缆接入网(HFC)等。

表 3.5 接入网传输系统分类

<table>
<tr><td rowspan="7">接入网</td><td rowspan="3">有线接入网</td><td>铜线接入网</td><td colspan="2">高比特率数字用户线(HDSL)、非对称数字用户线(ADSL)等</td></tr>
<tr><td>光纤接入网</td><td colspan="2">光纤到路边(FTTC)、光纤到户(FTTH)等</td></tr>
<tr><td>混合网</td><td colspan="2">混合光纤/同轴电缆接入网(HFC)等</td></tr>
<tr><td rowspan="3">无线接入网</td><td rowspan="2">固定无线接入网</td><td>微波</td><td>一点多址(DRMA)、固定无线接入(FWA)等</td></tr>
<tr><td>卫星</td><td>基小型天线地球站(VSAT)、直播天线等</td></tr>
<tr><td>移动接入网</td><td colspan="2">无绳电话、蜂窝移动电话、卫星通信等</td></tr>
<tr><td>综合接入网</td><td colspan="3">交互式数字图像(SDV)、有线+无线等</td></tr>
</table>

在无线接入方面,包括固定无线接入网和移动接入网,出现了微波一点多址、蜂窝以及微蜂窝技术、卫星通信等。

各种方式的具体实现技术多种多样,且各具特色,主要有以下几种技术措施:一是以原有铜质导线线路为主,在非加感的用户线上通过采用先进的数字信号处理技术来提高双绞铜线对的传输容量,向用户提供各种业务的接入手段;二是以光缆为主要传输,经同轴电缆分配给用户;三是全光化的实现,包括光纤到家庭等多种形式;四是以无线为主的接入方式。

3.4.2 基于协议的接入

1. 以主机方式通过 SILP/PPP 协议入网

以主机方式通过 SILP/PPP 协议入网如图 3.19 所示。客户通过使用 SILP/PPP 协议的通信软件,借助电话网和 CHINANET 提供的接入设备,可使客户自己的计算机随时成为 Internet 上的一台主机,接入全球 Internet,文件传送协议(FTP)需要配置完成。

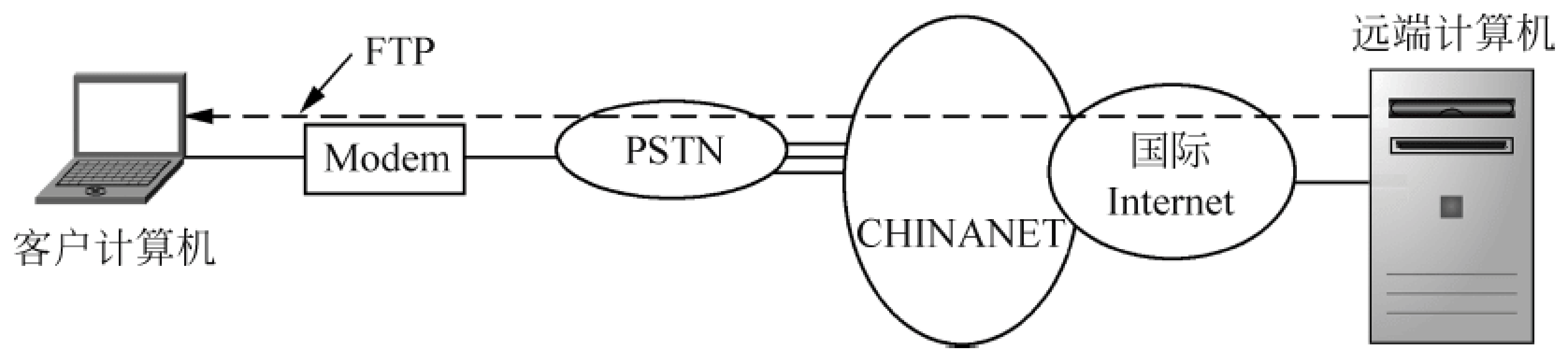


图 3.19　通过 SILP/PPP 协议入网示意图

2. 经分组网以 TCP/IP 协议入网

该方式类似于专线入网方式，所不同的是传输媒介是分组网的虚电路（SVC 或 PVC）而不是物理上的专线，客户在分组层上使用 TCP/IP 协议。该方式的接入示意图如图 3.20 所示。

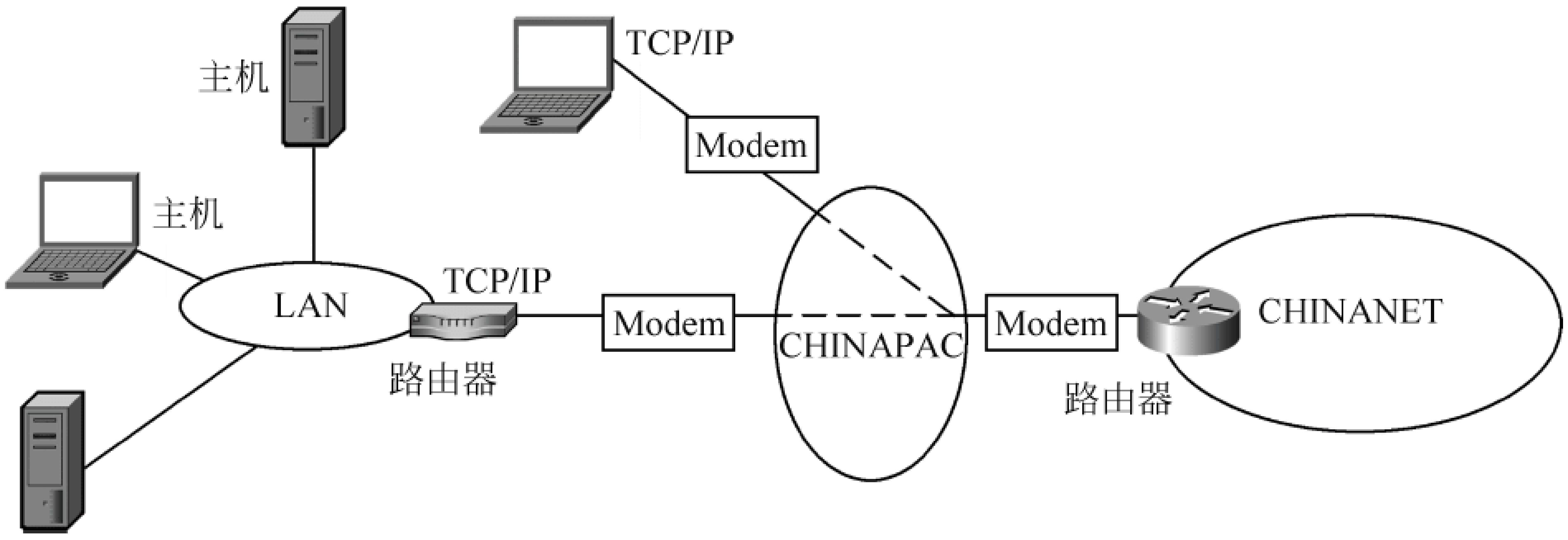


图 3.20　经分组网以 TCP/IP 入网

该接入方式对客户的要求，除通信软件需支持 X.25 协议外，其他等同于电话网主机入网方式或专线入网方式的要求，如需申请 IP 地址和主机名，考虑申请域名和支持相应的路由协议等。

3. 以帧中继方式入网

该方式通过帧中继网将客户接入 CHINANET，效果也类似于主机和专线上网。其特点是通信效率高，且租费较低。接入方式如图 3.21 所示。图 3.21 中路由器主要负责客户与外界的相互通信和路由控制。

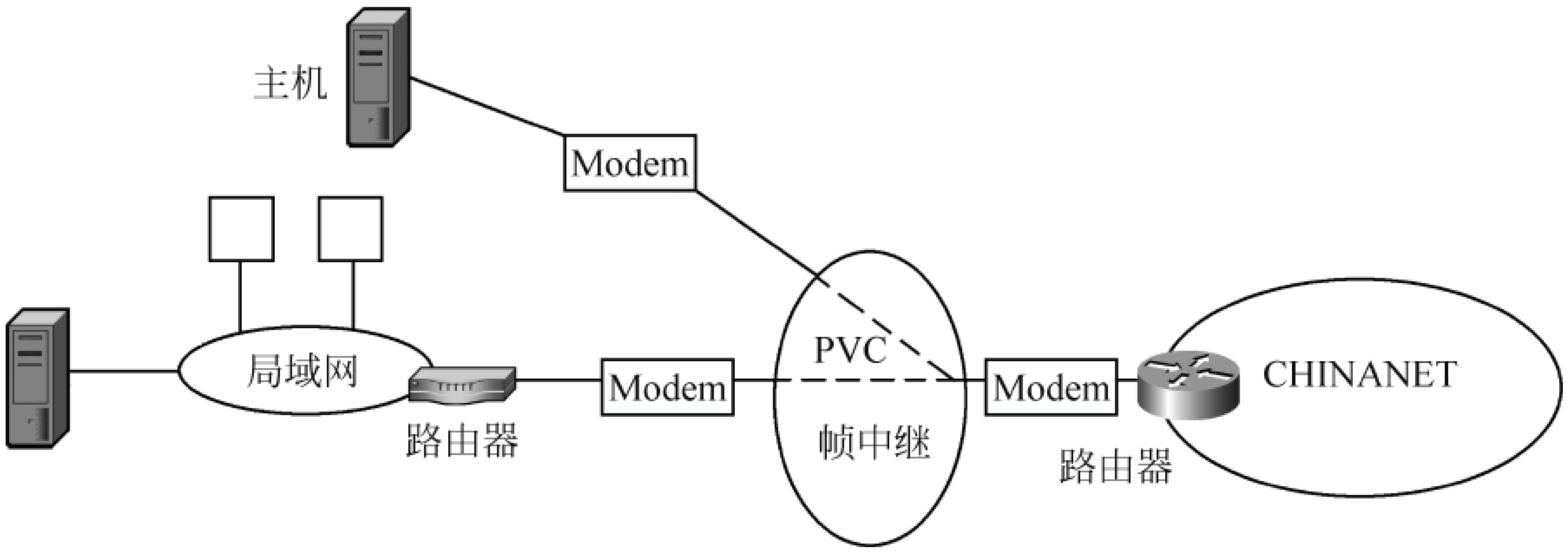


图 3.21　通过帧中继方式入网

客户端除需支持帧中继的通信软件以外，其余与主机或专线入网方式的要求相同，也需申请 IP 地址和主机名，考虑申请域名和确定路由协议，网络的数据传输速率可为 9600bps～

2.048Mbps。

3.4.3 以专线方式和电话拨号接入

1. 以专线方式入网

该方式主要是通过租用专线将客户接入 Internet。一般比较典型的方式为：用户将自己的相关计算机接到一个局域网上，再通过一个路由设备（通常为专用的路由器）经专线与 Internet 相连，接入方式如图 3.22 所示。

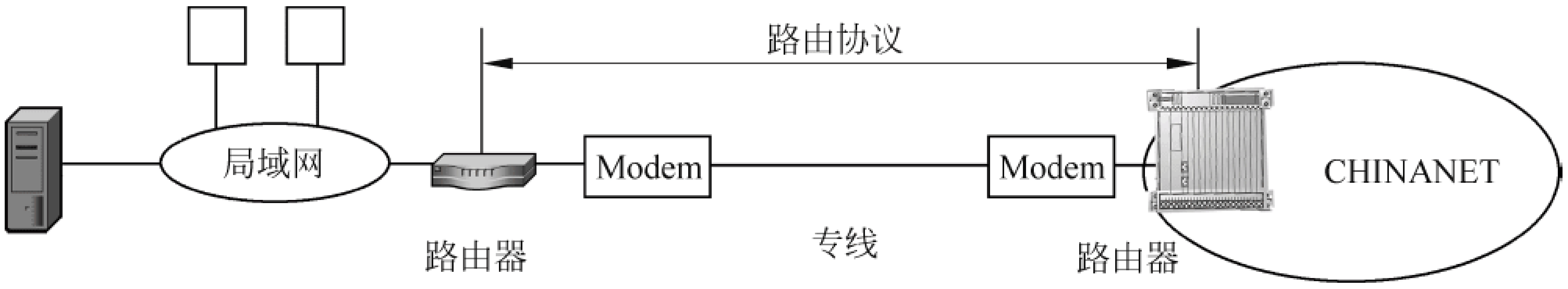


图 3.22 通过专线入网

专线方式主要适用于带有大量客户，需要传送大量信息及随时保持与 Internet 联系的情况。这种入网方式的费用相对较贵，但一旦建立连接，客户就把自己的计算机接入 Internet，可享受 Internet 上的所有应用并方便其他客户的访问，传输速率可达 1200bps～2.048Mbps 以上。

客户端的路由器要与 CHINANET 上的路由器连接，路由器之间需选择一定的协议，这个协议就称为路由协议。CHINANET 提供的支持客户接入的常用路由协议有 STATIC、OSPE、IS-IS、BGP 等。

2. 通过电话拨号上网接入

客户可以利用当前的电话网，随时方便地接入 CHINANET，CHINANET 为公众提供了一台接到 Internet 上的主机，客户在该主机上申请一个账号就可在每次通信时，先经由电话拨号登录到这台主机上，通过这台主机进入 Internet，而客户的计算机就相当于接入这台主机上的一个终端，其接入方式如图 3.23 所示。

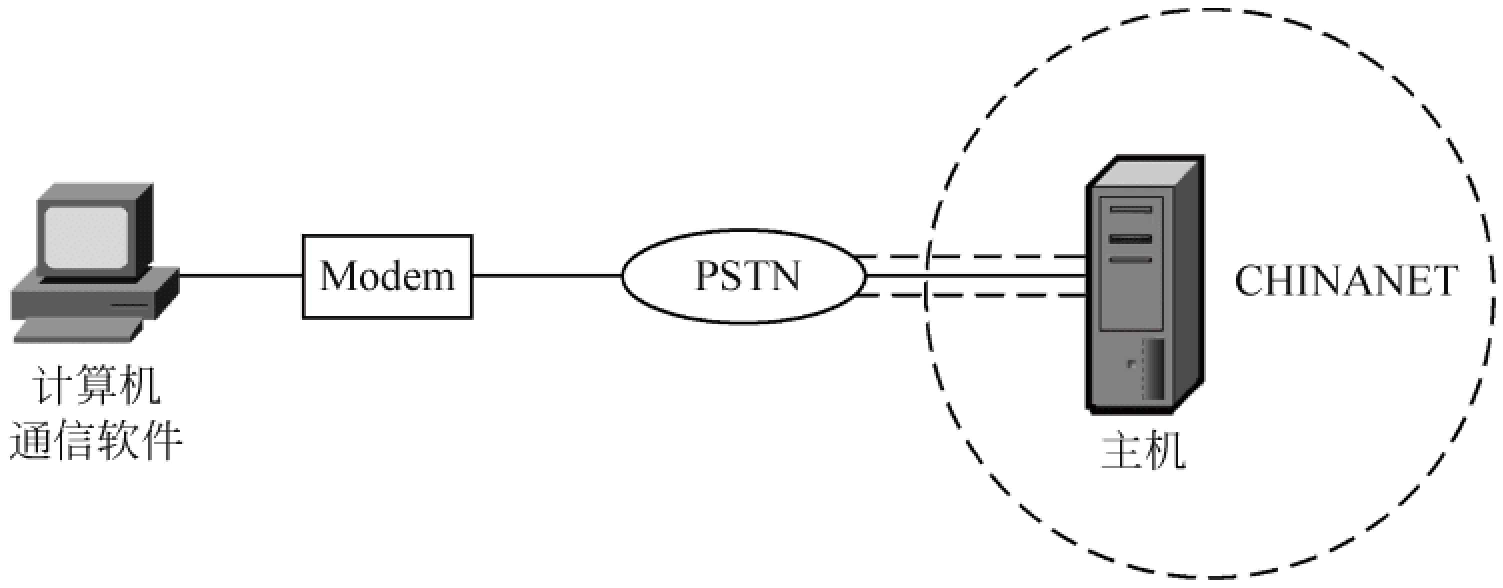


图 3.23 通过电话拨号接入方式

这种方式较为经济，客户可以通过接入主机，使用 Telnet、FTP、E-mail、Gopher 等各类应用，适用于业务量小的单位或个人。Modem 的传输速率可达 1200bps、2400bps 或 9600bps。因该方式为终端入网方式，故其使用 Internet 的某些应用还有不同，如要用 FTP 将文件从远端计算机传到客户端的计算机上的过程如图 3.24 所示，需分两步完成：

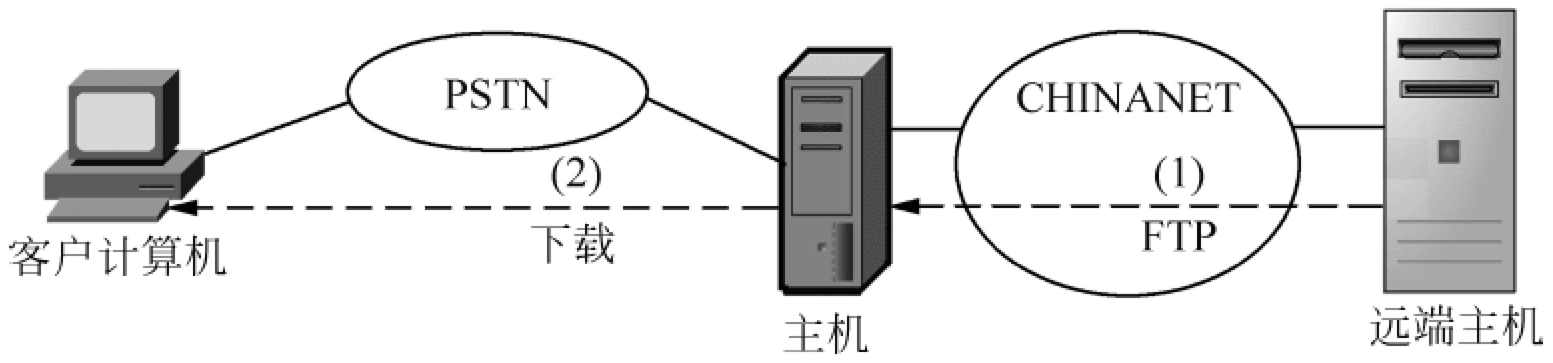


图 3.24 文件传到客户端

第一步，先把文件从远端计算机传到 CHINANET 的主机上。

第二步，用通信软件（如 Kermit）将文件传到客户端的计算机上。

3.4.4 ADSL 宽带接入

不对称数字用户线（Asymmetric Digital Subscriber Line，ADSL）是一种利用电话铜线的高速不对称用户环路传输技术。由于上行和下行的传输码率不相等，即所谓的不对称，避免了用户侧的近端干扰问题，从而得以提高了传输码率，延长了传输距离。在较多应用 ADSL 接入技术中，以下主要介绍其中的两种方式。

1. 基于计算机网络方式

ADSL 组网方式的思路主要来源于计算机网络，其拓扑结构如图 3.25 所示。ADSL 局端设备通过一局域网交换机连接在骨干网上，其接口为 10/100BT（Base-T 简写）以太网口，局端设备还具有集中功能，每一块用户卡支持 4 个 ADSL Modem；用户计算机通过 10BT 以太网接口接入网络，每个 ADSL Modem 具有网桥功能；整个接入系统通过局域网交换机的 10BT 接口进行网络管理，网络管理协议为简单网关协议（SNMP），可对整个接入网络进行设置、维护等方面的管理。

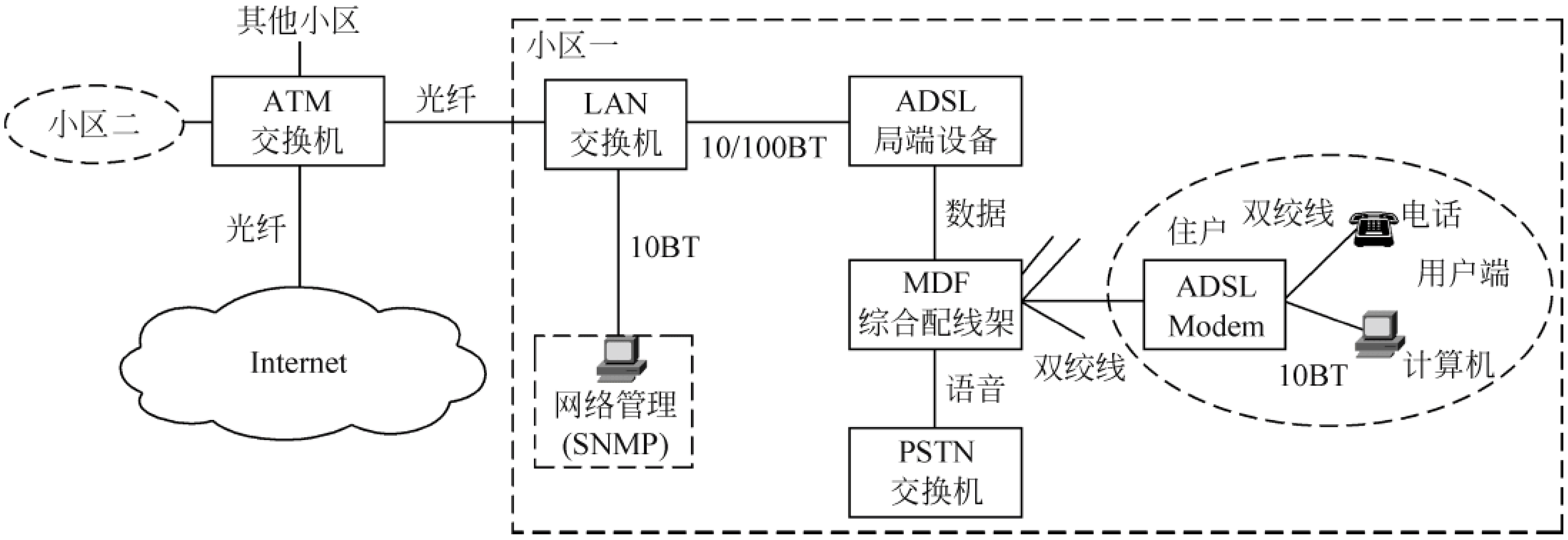


图 3.25 基于计算机网络方式的 ADSL

由于采用 CAP 线路编码和调制技术，在抗噪声干扰方面综合运用了纠错编码等技术，可使上下行最高速率达 1Mbps 和 7Mbps，当它运用于混合线路时，传输距离可达 3.8km。

2. 基于 ATM 信元传输的 ADSL 技术

这是一种端到端 ATM 信元传输的组网方式，即从局端到用户端的传输模式为 ATM 信元，如图 3.26 所示。与上一种方式不同的是，ATM 局端设备直接通过光纤连接到 ATM 骨干网上，局端设备具有集中功能，它的每一块用户卡可以支持多个用户，用户端计算机则通过 ATM 25.6Mbps 接口接入 ADSL Modem，整个 ATM 网络的线路编码和调制方式可

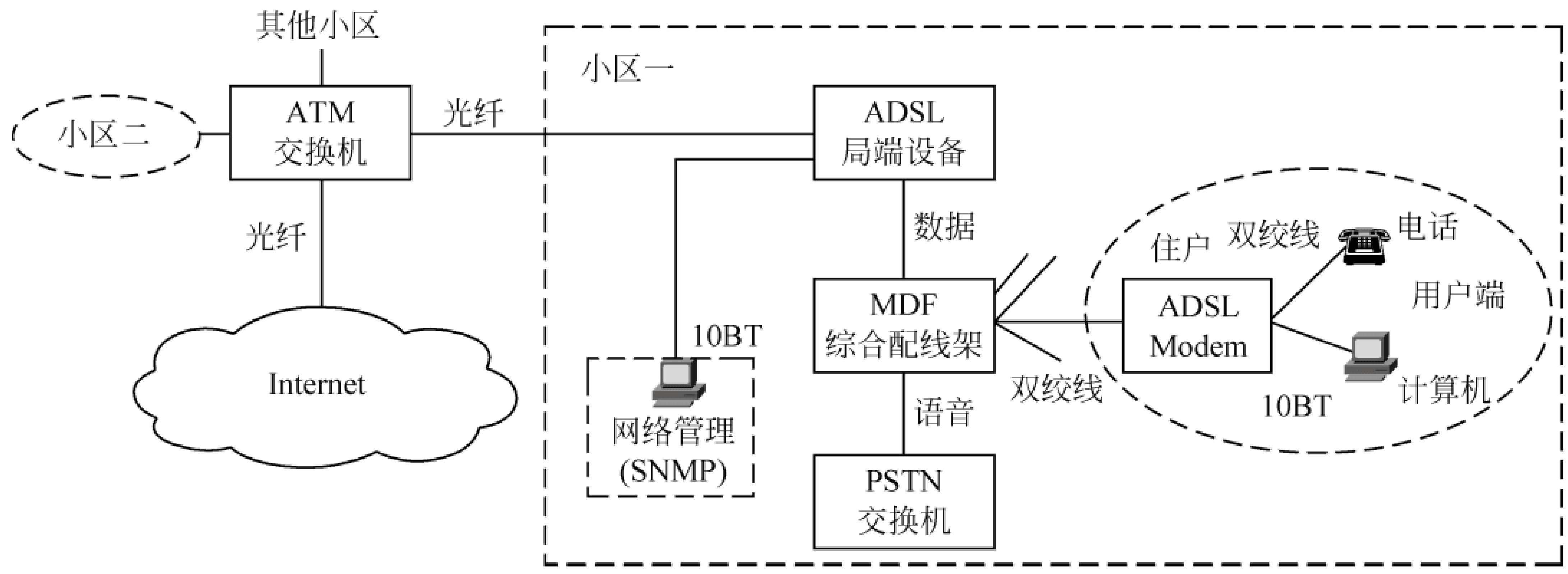


图 3.26 基于 ATM 信元光纤直接传输的 ADSL

以采用 DMT、无载波幅相调制(Carrierless Amplitude/Phase Modulation,CAP)等技术,使得下行速率得到有效提高。该系统通过 SNMP 协议对整个网络进行网络管理和维护。

3.4.5 混合网络和无源光网络

1. 混合网络(HFC)

由于接入网的应用环境复杂多变,采用单一的技术有时会难以满足不同用户的业务需求,因此混合接入网的技术方案也在实际应用中得到了发展。HFC(Hybrid Fiber Coaxial)是指光纤同轴电缆混合网,它是一种新型的宽带网络,采用光纤到服务区,而在进入用户的"最后 1 公里"采用同轴电缆。最常见的也就是有线电视网络,HFC 系统结构如图 3.27 所示。混合网简单归纳为以下几种。

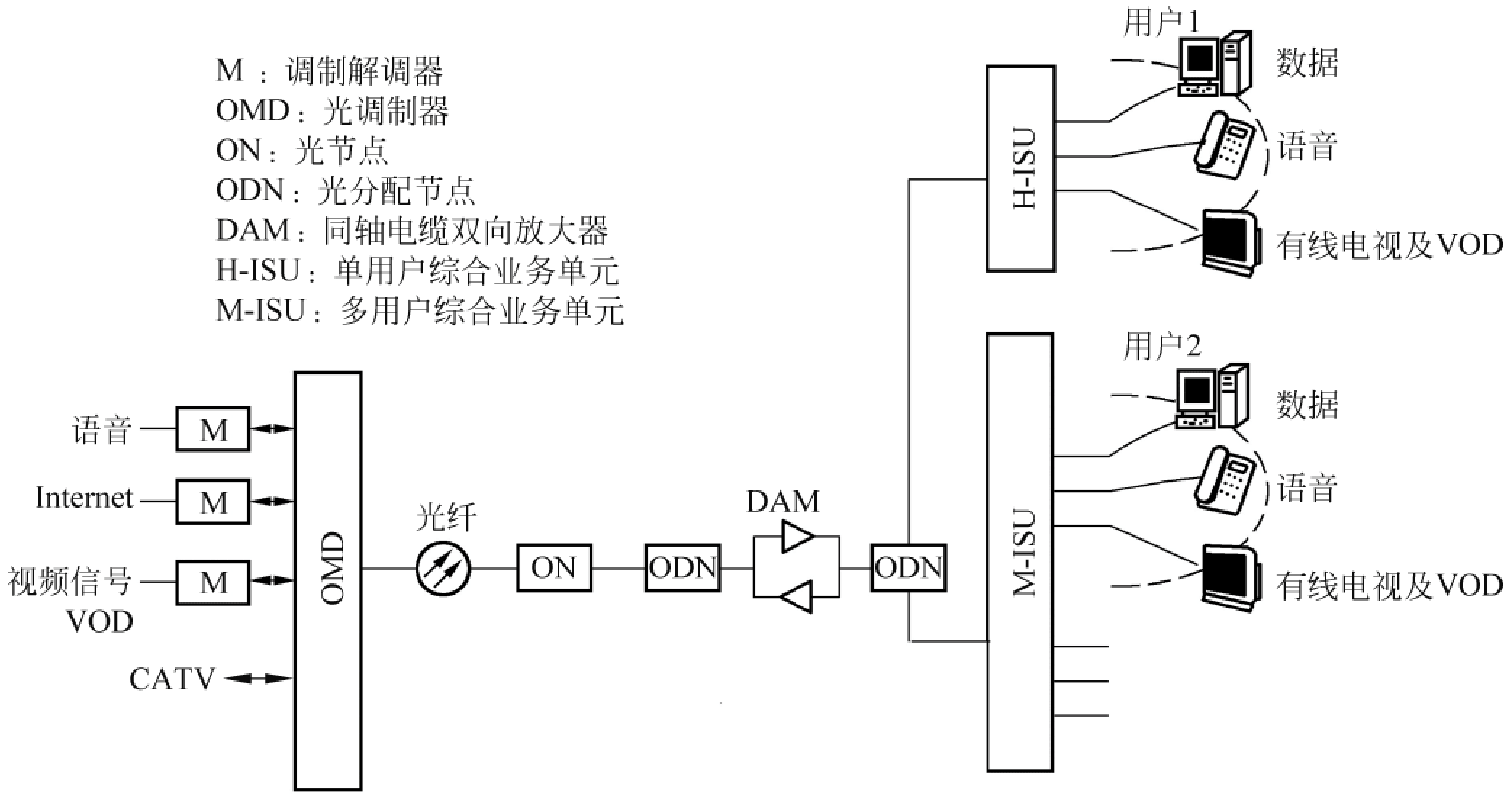


图 3.27 HFC 的系统结构示意图

(1) 窄带无源光网络(PON)+HFC 混合接入:这个混合接入方案的特点是充分利用 PON 双向多点的传输优势和 HFC 的单向分配型多点传输优势,实现优势互补。系统的基本结构是两套独立的基础设施,但可以通过 HFC 的光节点给 PON 的 ONU 供电。由于是

两套独立的基础设施，系统的建设比较灵活，可以先建 PON，以解决电话和数据的双向通信业务，以后再建 HFC 以满足 CATV 的需求。

(2) 数字环路载波(DLC)＋单向 HFC 混合接入：由于 DLC 在传输电话业务方面比 PON 要经济，尤其是采用标准中继接口和 V5 接口的 DLC 系统，其费用有望更低。但 DLC 系统的多点传输能力和业务的透明性不如 PON 系统，不是长期发展方向。

(3) 有线＋无线混合接入：有线与无线的混合接入也是一种优势互补的接入方案，其典型应用有 3 种：用无线代替有线的引入线部分，其他均为有线；用无线代替有线的配线和引入线部分，公共馈线仍为有线；用无线代替整个有线接入网，直接与本地交换机相连。

2. 无源光网络(PON)

以太网无源光网络(Ethernet Passive Optical Network，EPON)是一种新型的光纤接入网技术，它采用点到多点结构、无源光纤传输，在以太网之上提供多种业务。它在物理层采用了 PON 技术，在链路层使用以太网协议，利用 PON 的拓扑结构实现了以太网的接入。因此，它综合了 PON 技术和以太网技术的优点：成本低、带宽高、扩展性强、方便的管理等。

光纤用户网以 FTTH 为最终目标，采用 SDH 和建立光接入网是实现宽带业务的两大步骤。EPON 与 10G EPON 基于相同的标准体系，使用统一的运维模式和管理机制，用户可以根据带宽需求灵活选择 EPON、GPON 等，实现按需平滑升级。

3.4.6 综合业务接入网

这里讲的综合业务接入网是指基于统一传输平台下的网络融合。目前，运营商所能支持的业务种类繁多(参考图 3.28，它的左面是业务网，右面是接入网)，有的接入网在前面已作介绍，现大体归纳如下。

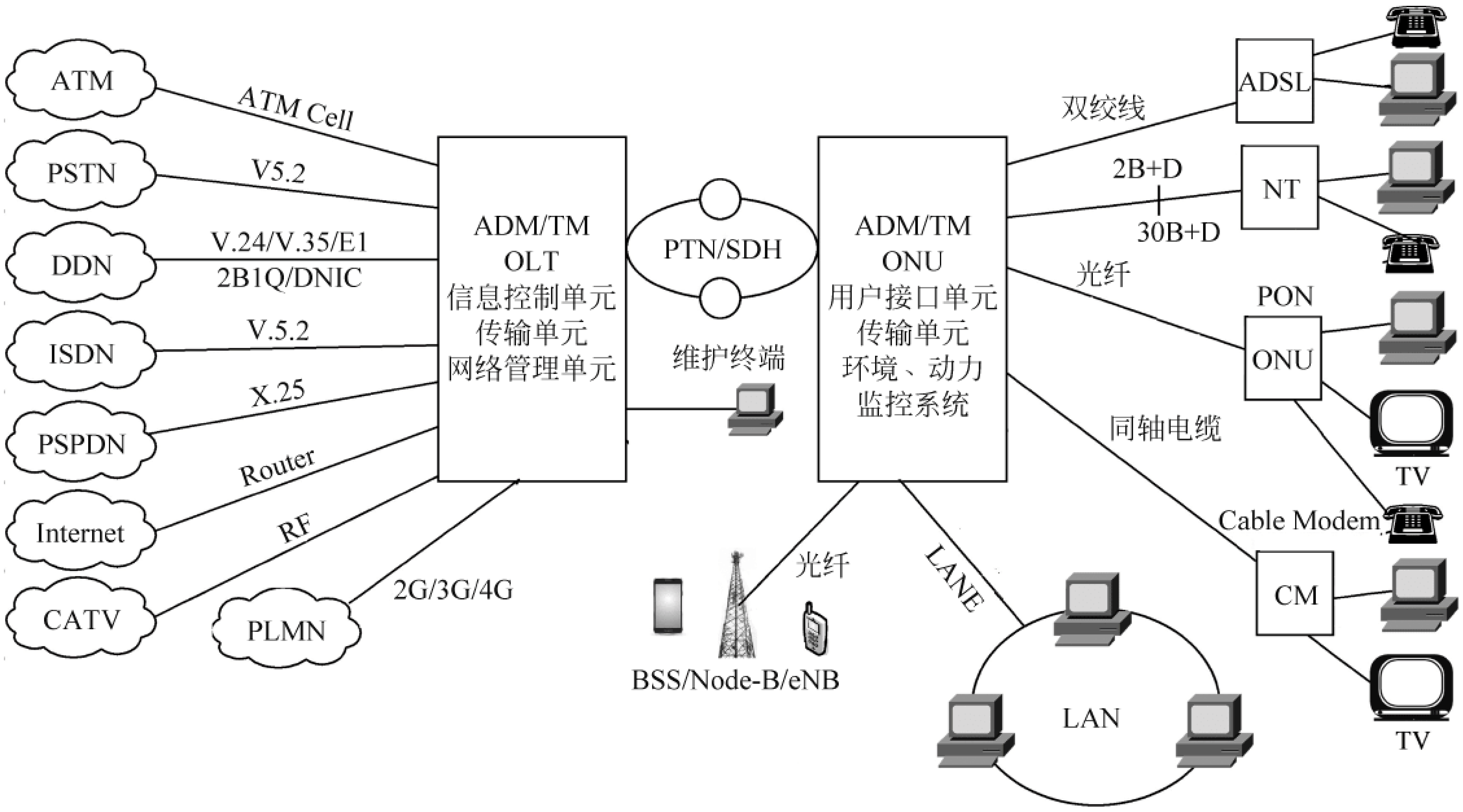


图 3.28 综合业务接入网

(1) 窄带 ISDN 业务：基本速率接入(BRA)和基群速率接入(PRA)，支持 ISDN 业务。

(2) 数字数据网(DDN)专线业务(V.24，V.35)：V.24 支持速率为 300Kbps、9600Kbps、19 200Kbps 等，支持同步和异步接口。V.35 支持速率为 $N\times 64$Kbps($N=1\sim 31$)。

(3) 数字租用线(Digital Leased Line，DLL)业务：支持成帧和不成帧的 2048Kbps 接口。

(4) 公交换分组网(PSPDN)业务：支持 POTS 用户以拨号或专线方式接入 PSPDN 网；支持 ISDN 用户以拨号或专线方式接入 PSPDN 网。

(5) Internet 接入业务：支持 POTS 用户以拨号方式接入 Internet(如 CHINANET，中国公众多媒体网等)；支持 ISDN 用户以拨号或专线方式接入 Internet；支持 DDN 用户以专线方式接入 Internet；支持动态链接库(dynamic link library)，使用户以专线方式接入 Internet。

(6) 公共陆地移动网(PLMN)业务：GPRS、3G、4G 接入 Internet，支持各种基于 IP 的业务。

(7) 2 线和 4 线音频专线业务：提供寻呼中心到无线发射机之间的通道，提供专线接入用户交换机的方式。

(8) CATV 业务：支持模拟方式同缆分纤传输。

(9) ADSL 业务：通过双绞线支持双向非对称宽带业务。

(10) LANE 业务：支持 10Base-T、100Base-Tx 接口，以及更高速率的接口。

(11) ATM UNI 接口：支持 ATM over SDH 等方式。

(12) 高速 Internet 接入业务：支持通过上述各种宽带接口接入 Internet。

(13) CATV 数字化传输业务：通过 ATM 接口接入 MPEG 2 编解码器，实现 CATV 数字化与宽带业务的一体化综合传输。

(14) PON 接入业务：通过 EPON、GPON 接入，实现 FTTH，为运营商接入各种业务。

习题

一、单项选择题

1. 能实现不同网络层协议转换功能的互联设备是(　　)。
   A. 集线器　　B. 交换机　　C. 路由器　　D. 网桥

2. 为什么路由器不能像交换机那样快地转发数据包？(　　)
   A. 路由器运行在 OSI 参考模型的网络层，因而要花费更多的时间来解析逻辑地址。
   B. 路由器中的路由表的建立时间比较长。

3. 在下列关于 Internet 的描述中正确的是(　　)。
   A. Internet 是一个协议。
   B. Internet 是 OSI 参考模型的上三层。
   C. Internet 是一个将许多网络互连在一起的网络。
   D. Internet 是 TCP/IP 协议栈。

4. 哪种设备的数据报转发时延最长？(　　)
   A. 网桥　　B. 路由器　　C. 交换机　　D. 网关

5. 使用(　　)能够将网络分割成多个 IP 子网。

A. 网桥　　B. 集线器　　C. 交换机　　D. 路由器

6. 下列不能直接用于网络互连设备的是(　　)。

A. 网关　　B. 路由器　　C. 交换机　　D. 三层交换机

二、是非判断题(将正确的题打上√)

1. 网络互连设备中,中继器工作于数据链路层。
2. 应用网关是在应用层实现网络互联的设备。
3. 中继器能起到延长传输距离的作用,对高层协议是透明的。
4. IPv4 中的 A 类地址是组播地址,其高位比特为 10。
5. ICMP 经常被认为是 IP 层的一个组成部分,它可用于传递 UDP 报文。
6. WWW 又称 3W 或 Web。
7. 文件传输协议(FTP)需要采用两个 TCP 连接来传输一个文件。
8. DHCP 是一个局域网的网络协议,传输层使用 UDP 协议。
9. SCTP 可提供面向连接的服务。
10. IP 提供不可靠的数据包传送服务,任何要求的可靠性必须由上层来提供。

三、简答题

1. 叙述综合业务接入网所能支持的业务种类。
2. 接入网是如何分类的?
3. 参考图 3.16,一个国外用户寻找一台叫 host.com.cn 的中国主机,说明其在 Internet 中通过域名的寻址过程。
4. CHINANET 提供哪些 Internet 接入?
5. 子网掩码为 255.255.255.0 代表什么意思?
6. 一个网络的现在掩码为 255.255.255.248,问该网络能够连接多少个主机?
7. 一个 B 类地址的子网掩码是 255.255.240.0。试问在其中每一个子网上的主机数最多是多少?
8. 一个 A 类网络的子网掩码为 255.255.0.255,它是否是一个有效的子网掩码?
9. 网络互连设备有哪些?它们各工作在 OSI 参考模型的哪一层?

四、计算题

已知地址块中的一个地址是 190.87.14.204/29。试求这个地址块中的最小地址和最大地址。地址掩码是什么?地址块中共有多少个地址?

第4章 数据通信网

数字数据网(Digital Data Network,DDN)是基于同步时分复用,采用物理层数字信道来传输数据的一种网络;数据分组交换网(CHINAPAC,也称PSPDN)采用X.25协议,广泛应用于广域网,处于OSI体系的下三层,IP数据包通过封装在X.25分组中传输;由于X.25网络体系不能很好地提供高速宽带的服务,因此,人们又在此基础上成功地开发出帧中继(FR),也称快速分组交换,它将数据流控制、差错检验和校正等交给由端到端操作的更高层协议去执行;属于宽带交换的异步传输模式(ATM)和IP曾经是两个互相竞争的技术,但最终实现了这两种技术的融合。本章主要介绍同属数据通信网范畴的DDN、X.25、FR、ATM相关协议及技术。

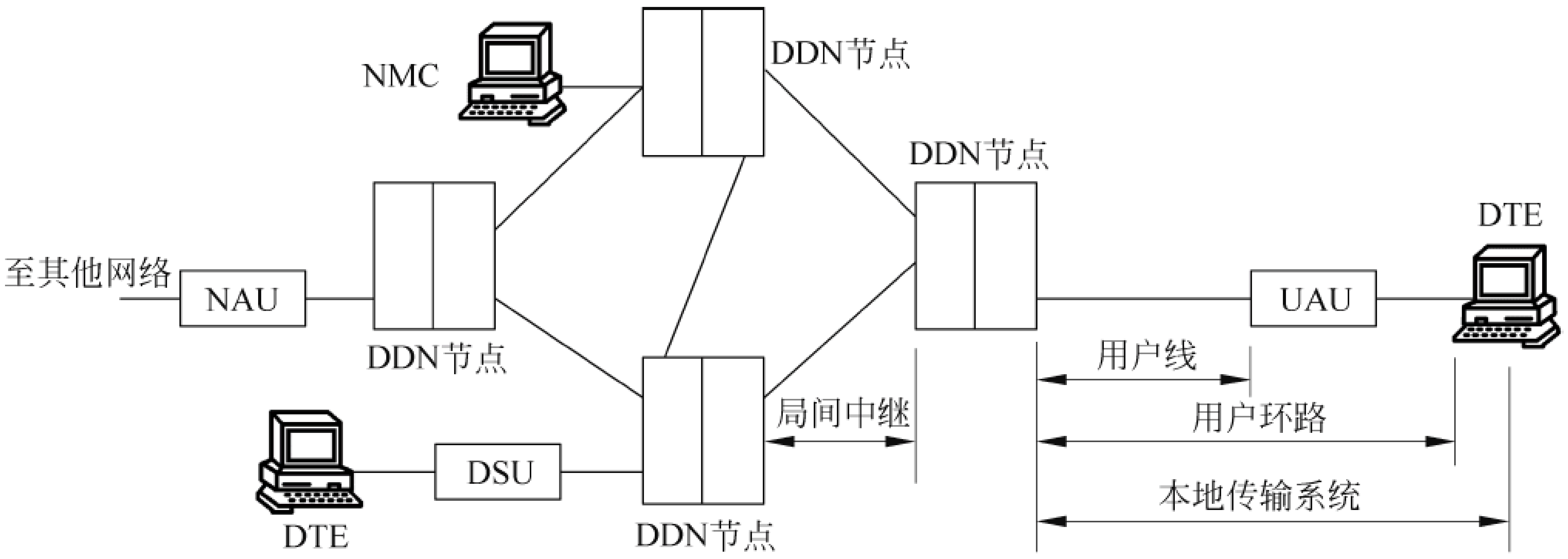
4.1 数字数据网

DDN为用户提供专用的数字数据传输信道,或提供将用户接入公用数据交换网的接入信道,也可以为公用数据交换网提供交换节点间用的数据传输信道。DDN可支持内外时钟或独立时钟方式,它是依附在电信传输网上的一个子网。

DDN具有以下特点:DDN是同步数据传输网,传输质量高;传输速率高,网络时延小;DDN为全透明网络,DDN任何规程都可以支持,满足数据、图像、语音等多种业务的需要;网络运行管理简便。

4.1.1 DDN组成

DDN组成如图4.1所示,它采用简单的交叉连接和复用装置,由本地传输系统、DDN节点(交叉/复用链接系统)、局间传输系统、网络同步系统和网络管理系统组成。



NAU:网络接入单元　　UAU:用户接入单元　　DSU:数据业务单元
DTE:数据终端设备　　NMC:网络维护管理中心

图4.1　DDN组成

1. 本地传输系统

本地传输系统,包含用户终端设备、用户线和用户接入单元,而用户线和用户接入单元又称用户环路。用户接入、用户环路及用户设备介绍如下。

1) 用户接入

用户接入是指用户设备经用户环路与节点连接的方式及业务种类。

用户接入可以是用户的网络接入单元(NAU),单个或成对出现,或是节点直接与终端相连接的接口;也可以是用户接入单元(UAU)或数据业务单元(DSU),如调制解调器或基带传输设备,以及时分复用、语音/数字复用等设备。DSU 属于 UAU 的一种。

2) 用户环路

用户环路是指用户终端至本地局之间的传输系统。用户是如何接入 DDN 的呢? 由于目前连接用户和 DDN 业务提供商的媒体主要是双绞线,用户接入主要采用 Modem 和 2B+D 线路终端设备,xDSL 设备也有所应用,PCM 是有条件大客户的另一选择。传输方式主要有 3 种。

(1) 四线全双工基带传输:适用于距离较短或传输速率较高的情况。

(2) 二线全双工基带传输:适用于距离 DDN 节点较远的用户。

(3) 模拟专线方式:作为前两种方式的补充,尽量少用。

3) 用户设备

用户设备包括用户终端和连接线。用户端设备可以是局域网,通过路由器连至对端,也可以是一般的异步终端或图像设备,以及传真机、电传机、电话机、数据终端、个人计算机、用户交换机、LAN 的桥接机、路由器等设备;连接线包括电话线、RS-232 电缆、RJ-25 芯插头及 LAN 使用的 10Base-T、10 Base-5、10 Base-2 等。

DTE 和 DTE 之间是全透明传输。

2. DDN 节点

DDN 是可为语音、数据、图像等信号提供半永久性连接电路的传输网络。所谓半永久性连接,是指 DDN 所提供的信道是非交换型的,用户之间的通信通常是固定的,可以在网络允许的情况下由管理人员或用户自己对传输速率、传输数据的目的地与传输路由进行修改,但这种修改不是经常性的,所以称作半永久性交叉连接或固定交叉连接。

DDN 节点包括时分复用器和数字交叉连接系统,主要完成接入、复用和交叉连接功能。通常在本系统中时分复用器是分级实现的:第一级,将来自多条用户线的信号形成 64Kbps 通信数据流;第二级,将 64Kbps 数据流按 32 路 PCM 系统格式进行时分复用。

数字接入和交叉连接系统(Digital Access and Cross-Connect System,DACS)用于通信线路的交接、调度管理。它的主要设备是智能化的数字交叉连接(DXC)设备、宽带管理器及供用户接入的设备。DXC 是一种具有交换功能的、智能化的传输节点设备,其实就是一个半永久性连接的、由计算机控制输入和输出数字流进行交叉连接的复用器和配线架。

DXC 设备系列代号为"DXC m/n",其中 m 表示输入数字流的最高复用等级,n 表示可以交换(或交叉连接)数字流的最低复用等级。m 的数字范围是 0~6,其含义如下:

(1) m=0,表示 64Kbps;

(2) m=1,表示 2Mbps(PDH)或 VC12(SDH);

(3) m=2,表示 8Mbps(PDH)或 VC-2(SDH);

(4) m=3,表示 34Mbps(PDH)或 VC-3(SDH);

(5) m=4,表示 140Mbps(PDH)或 155Mbps(SDH);

(6) m=5,表示 622Mbps(SDH);

(7) m=6,表示 2.5Gbps(SDH)。

DXC 的设备由同步电路、交换矩阵(交叉连接)和微机处理组成。其中,同步电路为网络提供精确的定时信号,用于进行时隙校准,DDN 的全系统时钟均由一个统一的同步时钟系统来提供;交换矩阵负责完成半永久电路的交叉连接;微机处理用于管理内部和外部操作,即维持操作系统,处理输入的命令和内部中断,执行时隙交接及监视系统等是否正常。

从组网功能上,DDN 节点可分为 2M 节点、接入节点、用户节点。

(1) 2M 节点:主要执行网络业务的转接功能。其主要有 2048Kbps 数字通道的接口;2048Kbps 数字通道的交叉连接;$N\times64$Kbps($N=1\sim31$)复用和交叉连接;帧中继业务的转接功能。因此通常认为 2M 节点主要提供 E1 接口;对于 $N\times64$Kbps 进行复用和交叉连接,起到收集来自不同方向的 $N\times64$Kbps 电路,并把它们归并到适当方向的 E1 输出的作用,或者直接对 E1 进行交叉连接。2M 节点如图 4.2 所示。

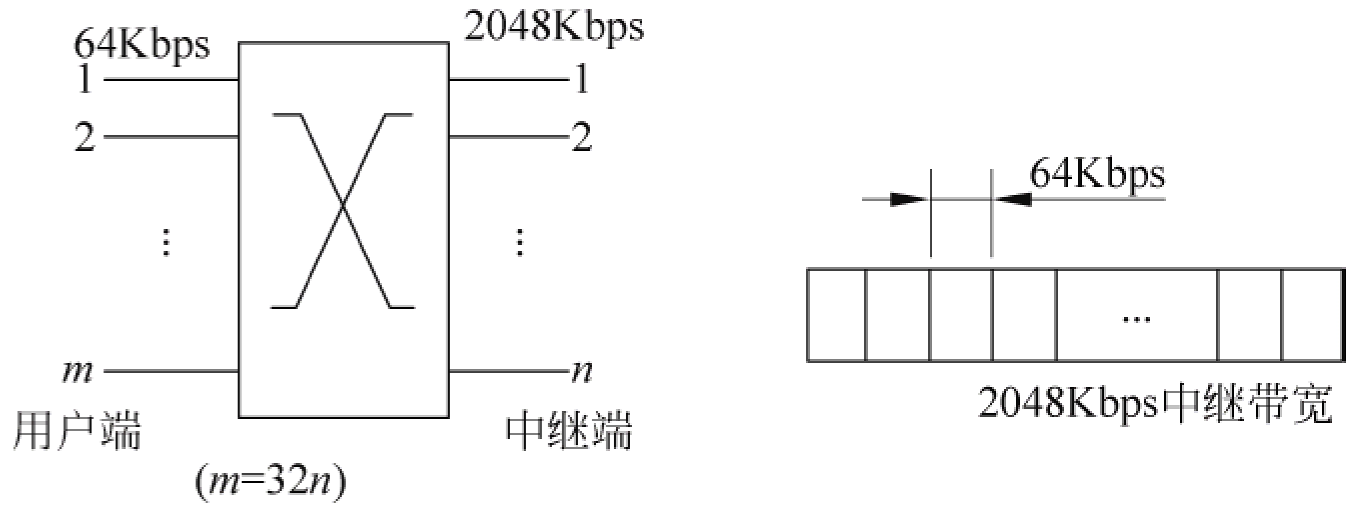


图 4.2　2M 节点

(2) 接入节点:主要为 DDN 各类业务提供接入功能。如 $N\times64$Kbps、2048Kbps 数字通道的接口;$N\times64$Kbps($N=1\sim31$)的复用;帧中继业务用户接入和本地帧中继功能等。

(3) 用户节点:主要为 DDN 用户入网提供接口,并进行必要的协议转换,包括小容量时分复用设备、LAN 通过帧中继互联的路由器等。用户节点也可以设置在用户处。

3. 局间传输系统

由光纤或数字微波通信系统组成的传输网是 DDN 的建设基础,PCM 高次群设备和光缆的大量使用及 SDH 光同步传送网的建设,使 DDN 具有以数字传输网作为网络建设基础的条件。DDN 常用的 TDM 复用,如 2M 数字电路帧结构,就是 PCM 的基群帧结构。

4. 网络同步系统

DDN 为了保证全网设备的同步工作,需要同步网络的支撑,它一般采用主从同步方式。

5. 网管中心

网管中心(NMC)采用分级管理,各级 NMC 之间能互换管理和控制信息,NMC 可以查看网络的运行情况、线路利用情况、统计报告、节点警告和故障报告等。各种情况要及时反映到 NMC,以便实现统一的网管功能。NMC 还可以方便地进行网络结构和业务的配置等。

4.1.2 DDN 用户接入方式及其应用

1. DDN 用户接入方式

1）模拟电话机接入方式

一种是普通的模拟电话机、传真机、电话交换机(PBX)等接入方式。另一种是话带(支持语音模式)Modem 接入方式，可以是二线的，也可以是四线(如 ISDN)的。

2）二线(或四线)基带接入方式

这类接入所用的设备通常是用户环路 Modem 或者其他终端，如 HTU(HDSL 终端单元)、VTU(VDSL 终端单元)等。这种传输方式采用回波抵消技术和差分二相编码技术。

3）基带传输＋TDM 复用接入方式

这类接入实际上是在上一种二线(或四线)基带传输的基础上，再加上 TDM 复用设备，为多个用户入网提供连接。

4）语音/数据复用接入方式

就是在现有的市话用户线上，采用频分或时分的方法实现电话/数据独立的数据复用传输。在 DOV(Data Over Voice)设备中，还可加上 TDM 复用，为多个用户提供入网连接。

5）采用 2B＋D 基本速率接口的 DTU(数据终端单元)接入

它通过线路终端(LT)设备与 DDN 节点连接，接入速率为 2B＋D 的 144Kbps(数据传输速率为 128Kbps)，采用二线全双工传输方式，可为多个用户提供入网。

6）PCM 数字线路接入方式

此接入方式类似于 E1、E3 等专用线路。当用户直接用光纤或数字微波高次群复用设备时，可与其他业务合用一套 PCM 设备，然后用其中的一路实现 2.048Kbps 的 DDN 接入。

7）通过 PCM 设备连接两个 DDN 节点的方式

当大量用户需要集中接入时，可以采用这种方式，将所传的数据信号复用到一条 2.048Kbps 的数字线路上，然后通过 PCM 的一个一次群信道连接到 DDN 上级节点上。

2. DDN 的应用

CHINADDN 以其优质的传输质量、智能化的网格处理以及灵活的组网方式，可以向用户提供多种业务，如图 4.3 所示。PVC(永久虚拟电路)是在源地址与目的地址之间的永久性硬件电路连接。SVC(交换虚拟电路)是根据实时交换要求建立的临时交换电路连接。两者的最大区别是：PVC 不论是否有数据传输，它都保持连接；而 SVC 在数据传输完成后就自动断开。以下介绍 DDN 提供的有关应用。

1）DDN 在计算机联网中的应用

DDN 作为计算机数据通信连网传输的基础，提供点对点、一点对多点的大容量信息传送通道。如利用全国 DDN 组成的海关、外贸系统网络。各省的海关、外贸中心首先通过省级 DDN，经过长途中继，到达国家 DDN 骨干核心节点。由国家网管中心按照各地所需通达的目的地分配路由，建立一个灵活的全国性海关外贸数据信息传输网络。

2）DDN 在金融业中的应用

DDN 不仅适用于气象、公安、铁路、医院等行业，也涉及证券业、银行、金卡工程等实时性较强的数据交换。

通过 DDN 将银行的自动提款机(ATM)连接到银行系统大型计算机主机。银行一般

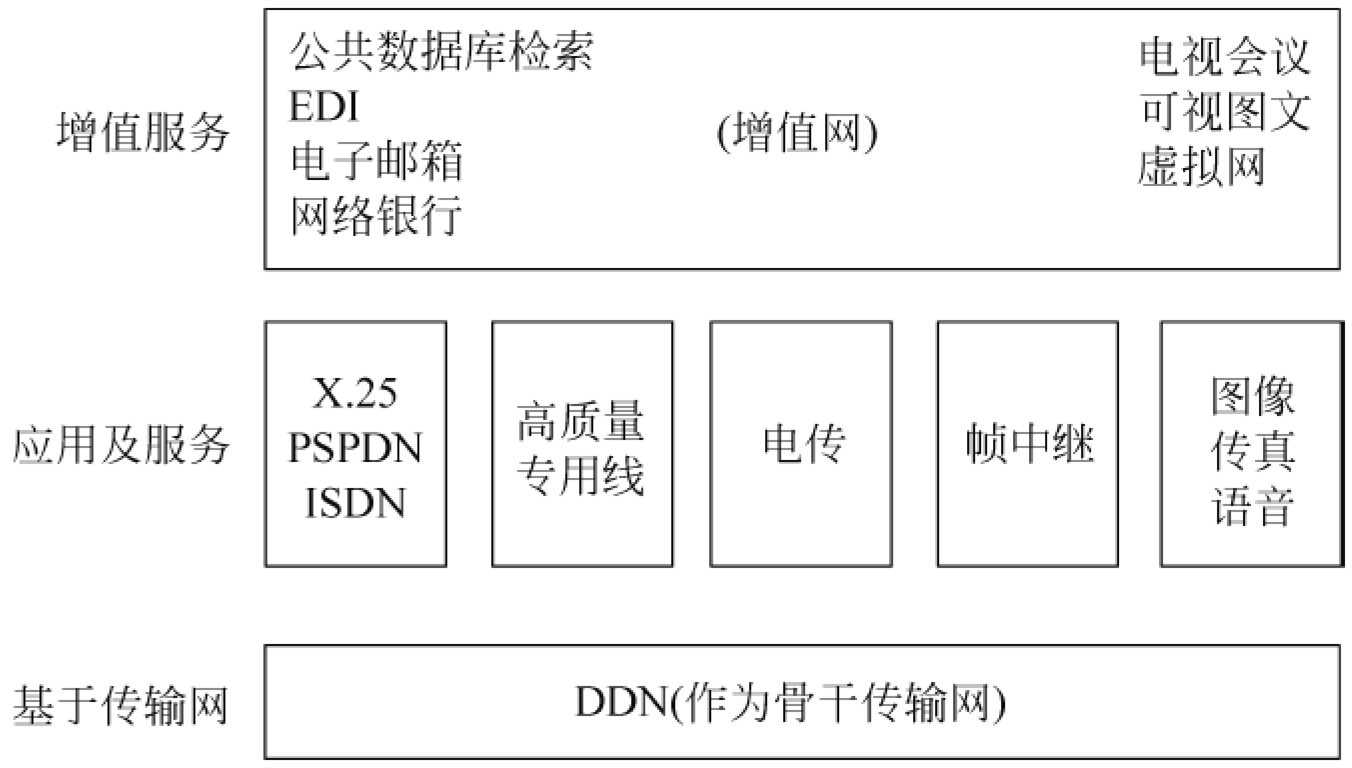


图 4.3 DDN 提供的业务

租用 64Kbps DDN 线路，把各个营业点的 ATM 机进行全市乃至全国连网。在用户提款时，对用户的身份验证、提取款额、余额查询等工作都是由银行主机来完成的。这样就形成一个可靠、高效的信息传输网络。

通过 DDN 发布证券行情，也是许多券商采取的方法。证券公司租用 DDN 专线与证券交易中心实行连网，大屏幕上的实时行情随着证券交易中心的证券行情变化而动态地改变，而远在异地的股民们也能在当地的证券公司同步操作，来决定自己的资金投向。

3）提供数据传输信道

目前，DDN 可为公用数据交换网、各种专用网、无线寻呼系统、可视图文系统、高速数据传真、会议电视、ISDN(2B＋D 信道或 30B＋D 信道)以及邮政储汇计算机网络等提供中继或用户数据信道。可为企业或办事处提供到其他国家或地区的租用专线。租用一条 DDN 国际专线，采用新的压缩技术，可以灵活地将 64Kbps 划分为 2.4Kbps(传送电报)、8Kbps(传送电话)、9.6Kbps(计算机连网)，且具有定时开放能力，即在约定的时间接通或拆除客户租用的数据电话，对客户而言更经济合理。

4）公用 DDN 的应用

DDN 可向用户提供速率在一定范围内可选的同步、异步传输或半固定连接端到端的数字数据信道。其中，同步传输速率为 600Kbps～64Kbps；异步传输速率用得较多的有 19.2Kbps、56Kbps 等；半固定连接是指其信道为非交换型，由网络管理人员在计算机上用命令对数字交叉连接设备进行操作。

5）网间连接的应用

DDN 可为帧中继、虚拟专用网、LAN 以及不同类型网络的互连提供网间连接；利用 DDN 实现大用户(如银行)局域网联网；可以使 DDN 平台成为一个多业务平台。

由于 DDN 独立于电话网，所以可使用 DDN 作为集中操作维护的传输手段。不论交换机 处于何种状态，它均能有效地将信息送到操作维护中心。

4.2 公共交换分组数据网

PSPDN 采用面向连接的虚(逻辑)电路交换方式。类似于电路交换方式，采用虚电路交换方式在通信前需要建立一条端到端的虚电路，通信结束后拆除这条虚电路。

PSPDN 采用 X.25 协议,其主要功能是描述如何在 DTE 和 DCE 之间建立虚电路、传输分组、拆除虚电路等,并且为用户提供了一些可选的业务功能和配置功能。

4.2.1 X.25 交换

1. X.25 网络的构成

X.25 是 DTE 与 DCE 进行点到点交互的规程,形成了 DTE 与 DTE 之间的通路。在一个 X.25 网络中,各实体之间的关系如图 4.4 所示。分组交换网的特点如下。

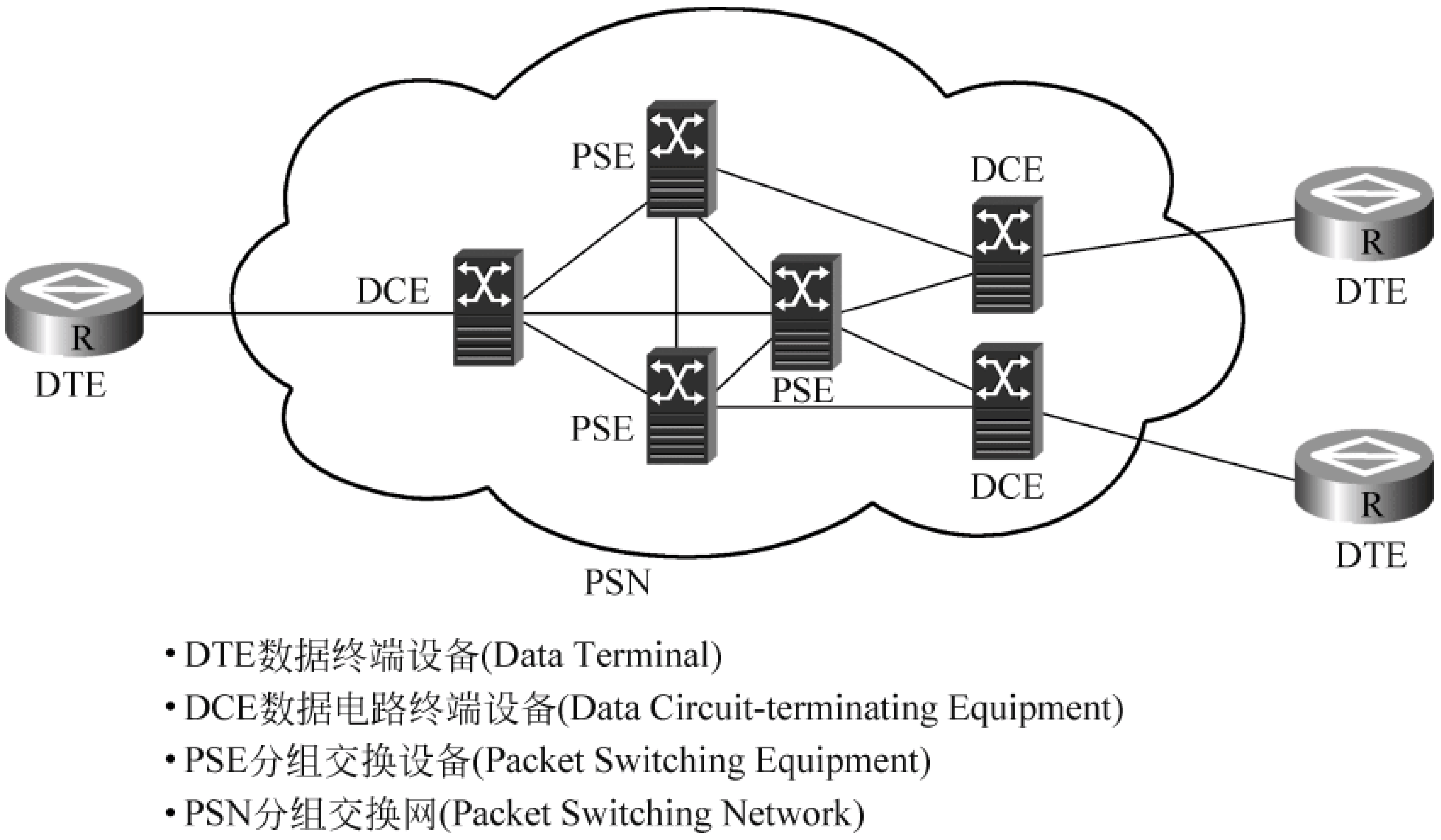


图 4.4 X.25 的网络的基本结构

(1) 分组交换具有多逻辑信道的能力,故中继线的电路利用率高。

(2) 可实现分组交换网上的不同码型、速率和规程之间的终端互通。

(3) 由于分组交换具有差错检测和纠正的能力,故电路传送的误码率极小。

(4) 分组交换的网络管理功能强。

2. X.25 各层对应格式

X.25 从第 1 级到第 3 级数据传送的单位分别是“位”“帧”和“分组”。当 DTE 向 DCE 传送信息时,第 2 级(链路层)接收到其上一级(分组层)的信息后,加上标志后通过下一级,就是物理层所提供的接口将信息传送出去,如图 4.5 所示。分组级以上的更高级都称用户级。

X.25 的物理层就像是一条输送信息的管道,它不执行重要的控制功能。控制功能主要由链路层和分组层来完成。物理层对 X.21 作了一些规定,如电路 T、R 一直处于工作状态,可以交换数据,C 和 I 也一直处于工作状态,其中 T 为发送分组,R 为接收分组,C 为控制信号,I 为指示位。X.21 实际上是 X.25 的分组流水线。

3. X.25 虚电路

虚电路是在主叫 DTE 和被叫 DTE 之间建立的一种逻辑连接,主叫或被叫的任何一方,在任何时候都可以通过这种连接发送和接收数据。

1) 虚电路和物理电路

虚电路是指在两台通信的 DTE 之间建立的连接,这种“电路”只在逻辑上存在,与电路交换中的物理电路有着质的区别。一旦在一对 DTE 之间建立一条虚电路,这条虚电路便

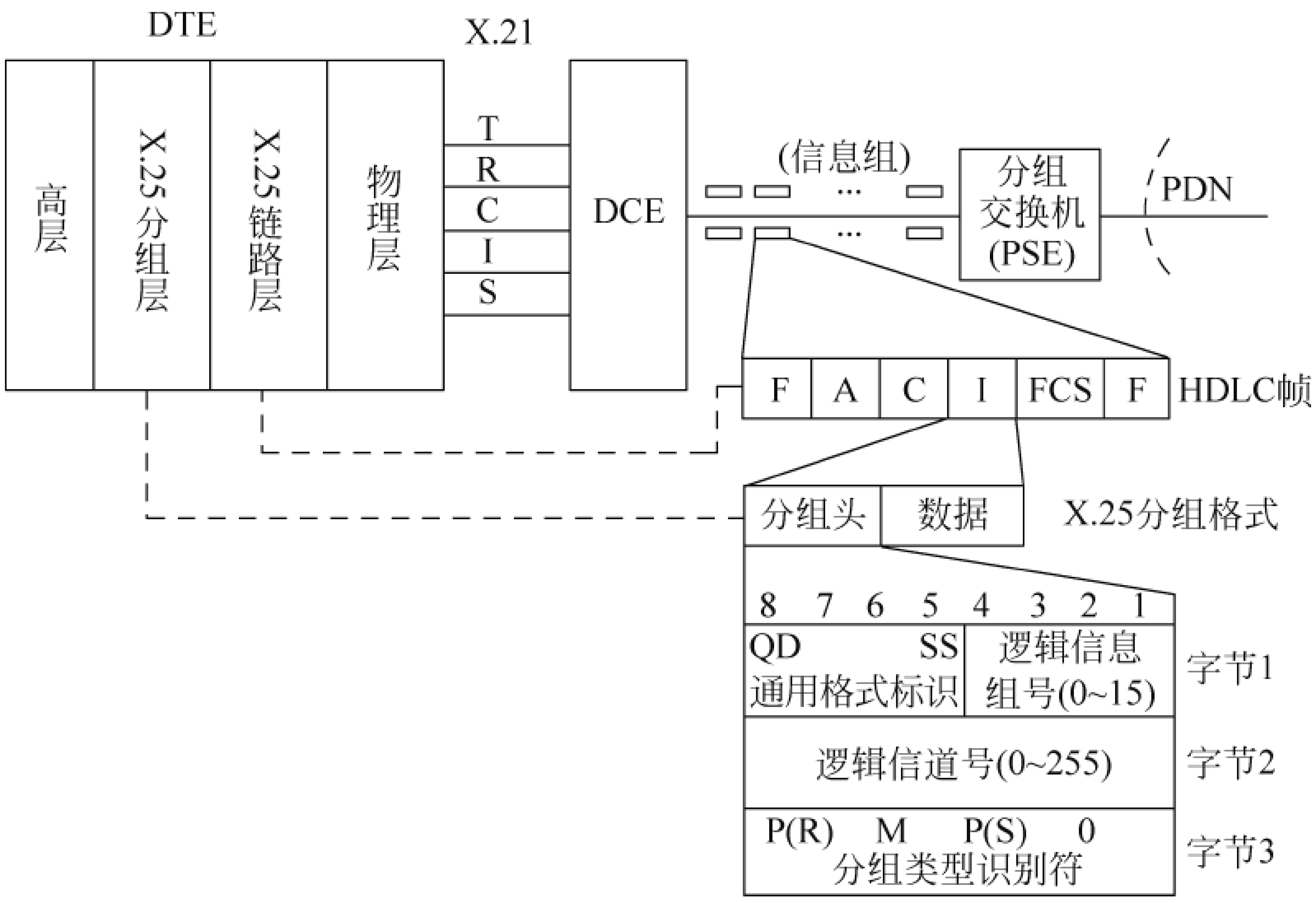


图 4.5　X.25 各层对应格式及关系

被赋予一个唯一的虚电路号，当其中的一台 DTE 要向另一台 DTE 发送一个分组时，它便给这个分组标上虚电路号交给 DCE 设备，DCE 就是根据分组所携带的这个号来决定如何在网络内部交换这个数据分组，使其正确到达目的地。X.25 第二层在 DTE 与 DCE 之间建立的一条链路被 X.25 第 3 层使用，最终呈现给用户的是可以使用的若干条虚电路。一个接口最多可以配置 4095 条虚电路，它采用的是统计时分复用方式。

一条虚电路并不独立占用线路和交换机的资源，也就是说在一条物理电路上可以同时开通多条虚电路，当某一条虚电路没有数据要传输时，线路的传输能力可以为其他虚电路服务。同样，交换机的处理能力也可以用于为其他的虚电路服务。因此，线路和交换机的资源能获得充分的利用。

2）虚电路和逻辑信道

虚电路是指在数据交换前，根据全网运行情况选择当时最佳传输路由，将传输数据的各段逻辑信道连接起来，组成一条完整的逻辑电路。这条逻辑电路只有在数据传输时才被分配占用，因而称作虚电路。

一条虚电路具有呼叫建立、数据传输和呼叫清除过程，永久虚电路可以在预约时由网络建立，也可以通过预约予以清除，而逻辑信道号是一种客观的存在，它有占用和空闲的区别，但是不会消失。

虚电路是主叫 DTE 和被叫 DTE 之间建立的虚连接，而逻辑信道是在 DTE-DCE 接口或网内中继线上分配的，代表子信道的一种编号资源，一条虚电路是由多个逻辑信道连接而成的。每一条线路的逻辑信道号的分配是独立进行的。

3）交换虚电路与永久虚电路

虚电路分为永久虚电路(Permanent Virtual Circuit，PVC)和交换虚电路(Switched Virtual Circuit，SVC)两种，顾名思义，PVC 用于两端之间频繁的、流量稳定的数据传输，而突发性的数据传输多用 SVC。

SVC 是指两个数据终端用户之间建立的临时逻辑连接。这种连接要通过数据终端的

拨号建立虚电路,它是一种双向式虚电路,既能用于发出呼叫,也能用于接入呼叫。在通信结束时,任一方都可发出拆线信号,释放虚电路。

PVC 是两个数据终端用户之间的虚电路永久连接,不需要拨号及拆线释放电路等过程,两端用户可以随时使用该逻辑信道。它适用于两个通信终端固定不变的通信需要,由于没有呼叫的建立,所以不存在呼叫冲突的问题。

4. X.25 协议分层结构与 IP 包传输

X.25 协议分为分组层、数据链路层、物理层 3 层,与 OSI 参考模型的下三层一一对应。对等层之间的通信通过对等层间的规程实现。

物理层定义了 DTE 和 DCE 之间的电气接口,以及建立物理的信息传输通路的过程。

数据链路层规定了在 DTE 和 DCE 之间的线路上交换帧的过程。从分层的观点来看,链路层好像是给 DTE 的分组层接口和 DCE 的分组层接口之间架设了一道桥梁,DTE 的分组层和 DCE 的分组层之间可以通过这座桥梁不断传送分组。链路层还进行帧的检错和恢复。

数据链路层采用平衡型链路访问规程 LAPB,采用了高级数据链路控制规程(HDLC)的帧结构,并且是它的一个子集。它通过置异步平衡方式(SABM)命令要求建立链路。LAPB 定义了 DTE-DCE 链路之间帧交换的过程及帧格式。虽然 LAPB 是作为 X.25 的第二层被定义的,但是,作为独立的链路层协议,它可以直接承载非 X.25 的上层协议进行数据传输。

支持 X.25 功能系列路由器可以直接使用 LAPB 协议进行串口封装,进行简单的本地数据传输;同时,支持 X.25 功能系列路由器的 X.25 还具备交换功能,也就是说,可以将路由器当作一台小型 X.25 分组交换机使用。IP 包通过 X.25 网络传送过程如图 4.6 所示,当局域网中产生的 IP 包传输到路由器后,路由器分析下一跳地址,决定通过某接口发送出去,这个接口封装了 X.25 协议。

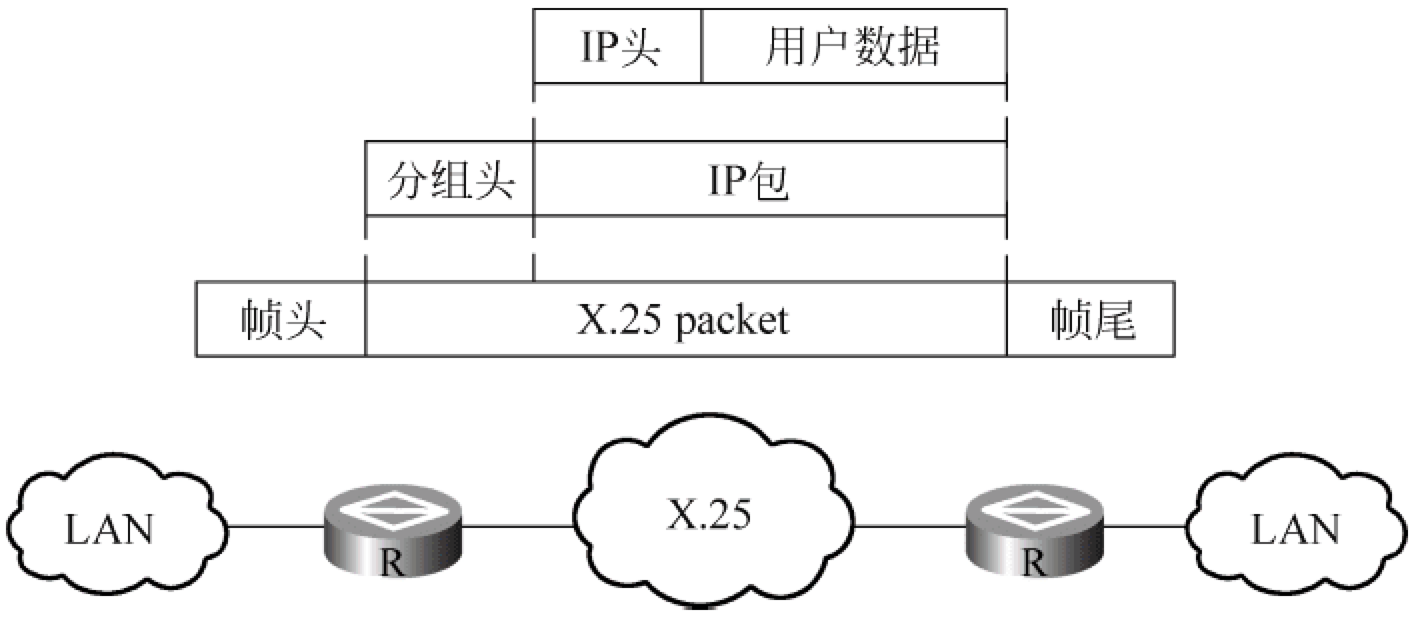


图 4.6 IP 包通过 X.25 网络传送过程

在路由器中 IP 包先传到路由器的分组层,分组层将 IP 包放在一个分组的数据段,在它前面加上分组头,然后下传给链路层。链路层看到的只是一个分组。链路层将分组当作帧的信息字段,加上帧头和帧尾封装成一个帧,而最终在物理链路上传送的是二进制的比特流。数据通过 X.25 网络传送到对端的路由器,路由器的各层协议将自己的结构层剥离,将数据送给上层协议处理。

4.2.2 用户接入分组网

用户终端设备按 X.25 协议接入 PSPDN，PSPDN 也可以按照 X.25 等相关协议实现与 PSTN、其他 PSPDN、ISDN、LAN 等网络的互连。

用户接入 PSPDN 如图 4.7 所示。用户终端设备接入方式有两类：一类是具有分组能力的分组型终端 P-DTE(简称 PT)，如计算机、智能终端等；另一类是以字符形式收发信息的一般终端，为非分组型终端 C-DTE(简称 NPT)，如异步字符终端、电话机等。

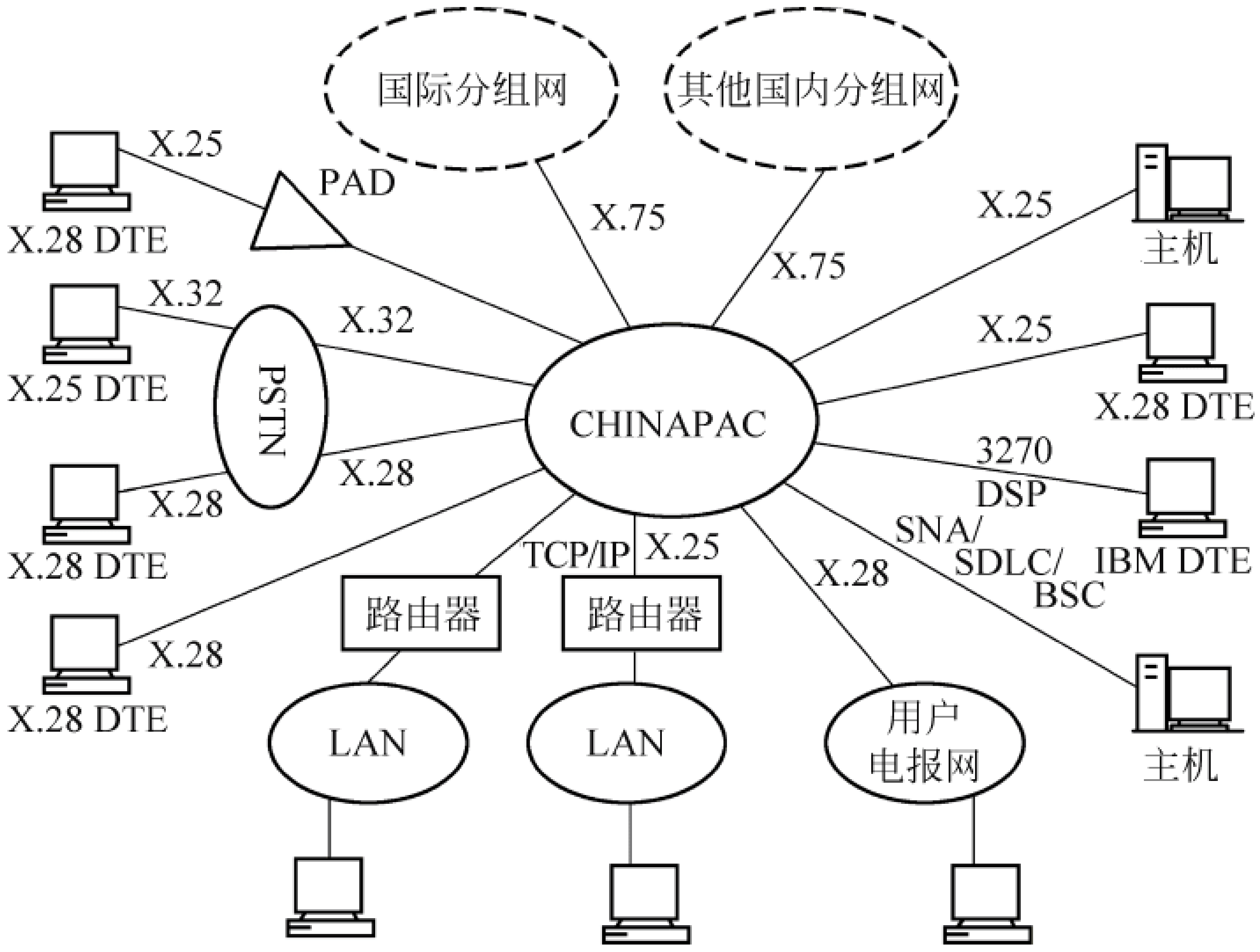


图 4.7 用户接入 PSPDN 示意图

终端设备(TE)进网有两种方式：一种是专线连接；另一种是电话线连接(通过 PSTN 网络拨号)。现在 NPT 可以用专线直接接入 PAD，也可以经 PSTN 与 PAD 连接。当 NPT 经 PSTN 接到 PAD 时，需先拨 PSTN 电话号码，在 NPT 与 PSTN 接口之间的通路建立好时再连接。

DTE 数据传输有两种方式：同步传输和异步传输。

通常计算机的 COM1、COM2 可收/发字符流。PAD 是具有分组交换设备的组成部分，也可以在用户端作为独立设备。

系统网络结构/同步数据链路控制(SNA/SDLC)终端采用 IBM 的 SNA/SDLC 规程的同步终端，可提供 2400Kbps、19.2Kbps、64Kbps 等速率的数据业务。

网络用户终端及入网方式可以归纳如下。

(1) 分组型终端及入网方式。

PT：必须具有分组形成能力，执行 OSI 参考模型的下三层，符合 X.25 接口规程功能，一般计算机上装有 X.25 网卡。通过相应的软件完成 PT 功能。PT 通过 PSTN 拨号接入网络服务应遵循 X.32 建议。

专线接入：使用 X.25 接口规程，可以提供 2400bps、4800bps、64Kbps 等速率的数据通信业务。

PSTN 接入：使用 X.32 接口协议，可提供 64Kbps、48Kbps、19.2Kbps、2400Kbps、4800Kbps、

9600Kbps 等速率的数据通信业务。

(2) 非分组型终端及入网方式。

NPT：只能收/发字符流，不具备分组拆/装成字符流的能力，必须通过分组拆/装设备(PAD)入网。

PSTN/专用网接入：使用 X.28 接口协议，可提供 300bps、1200bps 等速率的数据通信业务。

电报网接入：使用 TEL/X.28 接口规程，可提供 50bps 速率的数据通信业务。

4.3 帧中继

帧中继(Frame Relay，FR)是 X.25 在新的传输系统、新型终端设备迅速发展的条件下，为适应急剧增长的 LAN 互连的形势而发展起来的技术，它只完成 OSI 的物理层和数据链路层核心层的功能，保存了 X.25 的链路层 HDLC 的帧格式，但不采用 LAPB 规程，而是按照 ISDN 标准，使用“D 信道链路接入协议”，在链路层实现链路的复用和转接，完全不用网络层，故得名帧中继。

4.3.1 FR 交换

1. FR 的特点

帧中继可以作为一种新型的数据传输网络，为了满足局域网的互连所需的大容量的传送，也为了满足用户对数据传输延迟小的要求。帧中继是基于虚电路(virtual circuit)的。由于帧中继数据单元至少可以达 1600 字节，所以帧中继协议十分适合在广域网中连接局域网。用户的路由器封装帧中继协议，作为 DTE 设备连接到帧中继网中的 DCE 设备，也就是帧中继交换机。

帧中继网与局域网的连接如图 4.8 所示，其中数据链路连接标识(Data Link Connection Identifier，DLCI)用于标识每一个 PVC，帧中继网络中的每一个连接都使用 DLCI 来标识。通过帧中地址字段的 DLCI，可以区分出该帧属于哪一条虚电路。

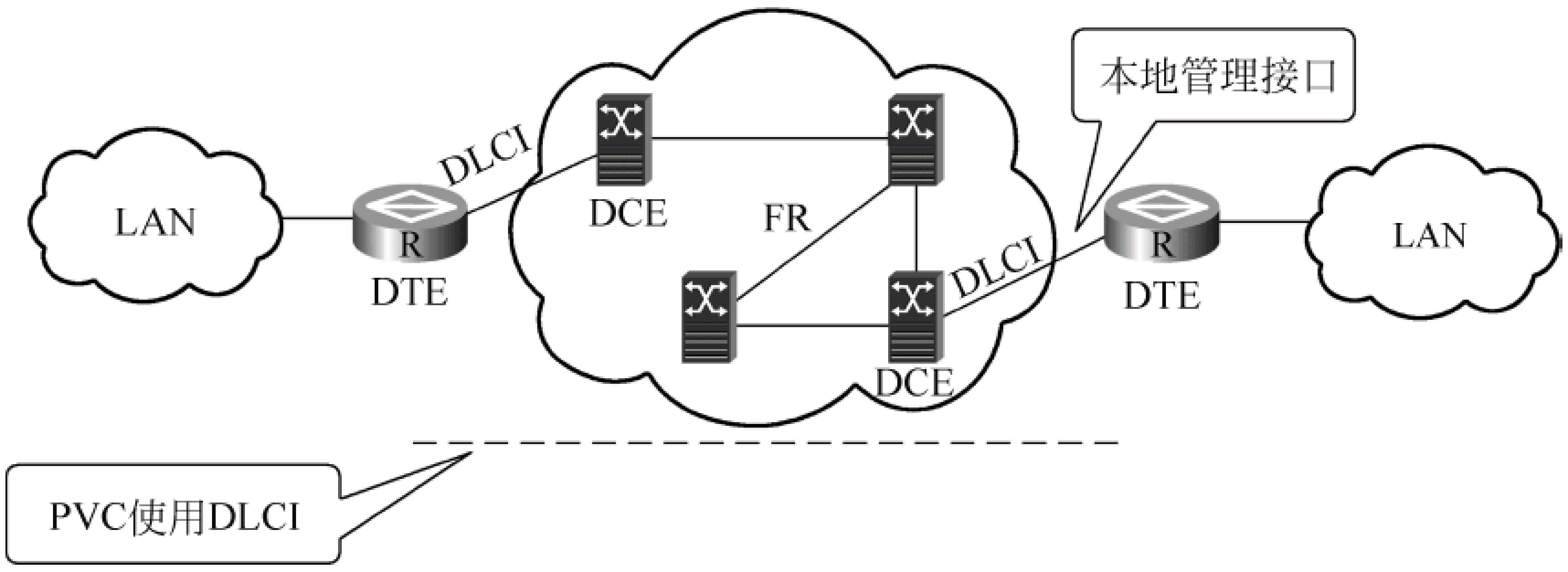


图 4.8 帧中继网与局域网的连接

本地管理接口(Local Management Interface，LMI)协议用于建立和维护路由器和交换机之间的连接。LMI 协议还用于维护虚电路，包括虚电路的建立、删除和状态改变。对于 DTE 侧设备，永久虚电路的状态完全由 DCE 侧设备决定。对于 DCE 侧设备，永久虚电路

的状态由网络来决定。在两台网络设备直接连接的情况下,DCE 侧设备的虚电路状态是由设备管理员来设置的。

帧中继是分组交换的改进,提高了处理速度,可概括为以下几个特点。

(1) 用于传送数据业务,要求传输速率高,信息传输的突发性大,各类 LAN 通信规程的包容性好。

(2) 帧中继采用虚电路技术,只有当用户准备好数据时才把所需的带宽分配给指定的虚电路,而且带宽在网络中是按照每一分组以动态方式进行分配的,因而适合于突发性业务的使用。传输链路是逻辑连接,而不是物理连接,在一个物理连接上可以复用多个逻辑连接,可以实现带宽的复用和动态分配。

(3) 简化了 X.25 的第 3 层协议,帧中继只使用了物理层和链路层的一部分来执行它的交换功能;利用用户信息和控制信令分离的 D 信道连接实施以帧为单位的信息传送,简化了中间节点的处理。帧中继采用了可靠的 LAPD 协议(LAPD 是 ISDN 的 D 信道链路层的协议),将流量控制、纠错功能留给智能终端去完成,简化了处理过程,提高了效率。

(4) 在链路层完成统计复用、透明传输和错误监测(不重复传输)。

(5) 用户传输速率一般为 64Kbps～2Mbps,根据用户需要,有的速率可为 9.6Kbps。较高为 8～10Mbps,以后将达到 34～45Mbps。

(6) 交换单元(帧)的信息长度比分组交换长,达到 1024～4096 字节/帧,预约的最大帧长度至少要达到 1600 字节/帧。因而其吞吐量非常高,其所提供的速率大于 2.048Mbps。

(7) 有合理管理带宽的机制,用户除实现预约带宽外还允许突发数据预定的带宽。提供一套合理的带宽管理和防止拥塞的机制,用户有效地利用预约的带宽,即承诺的信息速率(CIR),还允许用户的突发数据占用未预定的带宽,以提高网络资源的利用率。

(8) 帧中继使用统计复用技术(即按需分配宽带),向用户提供共享的网络资源,每一条线路和网络端口都可以由多个终点按信息流共享,提高了网络资源的利用率。

(9) 与分组交换一样,采用面向连接的交换形式,可提供 SVC(交换虚电路)业务和 PVC(永久虚电路)业务。

2. FR 帧结构

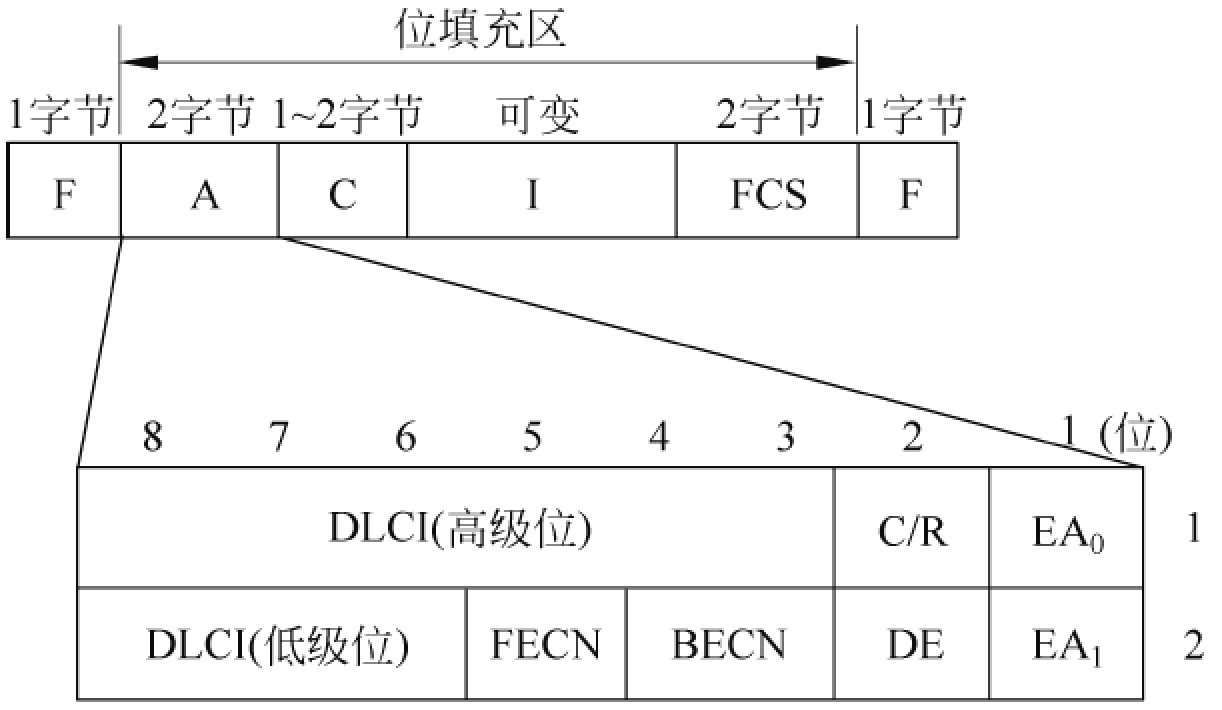
帧中继网络只进行 CRC 校验,丢弃出错的帧,完全的差错控制和重发留给终端去解决。

由于许多数据通信协议都普遍符合 LAPD(Line Access Procedure on the D-channel)链路层协议,LAPD 子协议就是用于帧中继 DCE 与 DTE 间相互工作的信令。LAPD 链路层具有如下内容:数据报协议,每个帧含有目的地址;每个源节点可以同时与多个目的节点通信。帧中继没有采取存储转发功能,因而具有与快速分组交换相同的优点。如图 4.9 所示就是帧中继的帧结构,以下对其进行说明。

(1) F:标志位,以 8 位组 01111110 表示,表示一帧的开始和结束,帧结构中其余部分为位填充区。

(2) A:地址字段,该字段用于区别同一通路上的多个数据链路的连接,以便实现帧的复用/分路。其长度通常是 2 字节,最大可以扩展到 4 字节,包括字段扩展位 EA、命令/响应指示位 C/R、帧可丢失位 DE、前向显式拥塞位 FECN、数据链路连接标识符 DLCI 和 DLCI 扩展/控制指示位 D/C 共 7 个组成部分。

(3) C:控制字段。LAPF 定义了 3 种类型的帧:



F：标志 A：地址 C：控制 I：信息 FCS：帧检验序列
DLCI：数据链路标识符 FECN：前向拥塞指示
BECN：后向拥塞指示 C/R：命令/响应指示
DE：可抛弃标志 EA：字段扩展位

图 4.9　帧中继的帧结构及地址字段格式

信息帧(I 帧)用来传送用户数据，但在传送用户数据的过程中，可以携带流量控制和差错控制，并且 LAPF 允许 I 帧使用 F 位；

监视帧(S 帧)专门用来传送控制信息，当流量控制和差错控制不能搭乘 I 帧时，就用 S 帧来传送；

未编号帧(U 帧)用来传送控制信息和安排非确认方式传送用户数据。

(4) I：信息字段。由整数倍的字节组成，包含的是用户数据位序列。

(5) FCS：帧校验序列字段(FCS)。能检测出任何位置上 3 位以内的错误、所有奇数个错误、16 位之内的连续错误和量的突发性错误。

3. FR 带宽控制参数

帧中继是统计复用协议，实现了带宽资源的动态分配，它通过为用户分配带宽控制参数，对每条虚电路上传送的用户信息进行监视和控制。帧中继网络为每个帧中继用户分配 3 个带宽控制参数：Bc、Be 和 CIR。同时，每隔 Tc 时间间隔对虚电路上的数据流量进行监视和控制。承诺信息速率(Committed Information Rate，CIR)是网络与用户约定的用户信息传送速率，即承诺信息速率。如果用户以小于等于 CIR 的速率传送信息，应保证这部分信息的传送。Bc 是网络允许用户以 CIR 速率在 Tc 时间间隔传送的数据量，即 Tc＝Bc/CIR。Be 是网络允许用户在 Tc 时间间隔内传送的超过 Bc 的数据量。

在 Tc 内，网络对每条虚电路进行带宽控制，可采用策略是：当用户数据传送量小于、等于 Bc 时，继续传送收到的帧；当用户数据传送量大于 Bc，但小于、等于 Bc＋Be 时，将 Be 范围内传送的帧的 DE 位置“1”，若网络未发生严重拥塞，则继续传送，否则将这些帧丢弃；当 Tc 内用户数据传送量大于 Bc＋Be 时，将超过范围的帧丢弃。

例如，如果约定一条 PVC 的 CIR＝64Kbps，Bc＝64Kbit，Be＝32Kbit，则 Tc＝Bc/CIR＝1s。在这一段时间内，用户可以传送的突发数据量可达到 Bc＋Be＝96Kbit，传送数据的平均速率为 96Kbps，其中，正常情况下，Bc 范围内的 64Kbit 的帧在拥塞情况下，这些帧也会被送达终点用户，若发生了严重拥塞，这些帧会被丢弃。

Be 范围内，即 32Kbit 帧的 DE 位被置为“1”，在无拥塞的情况下，这些帧会被送达终点

用户,若发生拥塞,则这些帧会被丢弃。当转发队列中的报文长度超过一个阈值,可以认为发生了拥塞。当拥塞发生,在该队列中的报文的 FECN 位将被置位。如果拥塞持续下去,相反方向的报文的 BECN 位将被置位。

4. 帧中继 DLCI 和地址映射

1) 帧中继 DLCI 分配

由于帧中继虚电路是面向连接的,本地不同的 DLCI 连接到不同对端设备,所以也可以认为本地 DLCI 就是对端设备的帧中继目的“逻辑地址”。

帧中继在单一物理传输线路上能够提供多条虚电路。通过帧中继帧中地址字段的 DLCI,就可以区分出该帧属于哪一条虚电路。DLCI 只在本地接口和与之直接相连的对端接口有效,不具有全局有效性,即在帧中继网络中,不同物理接口上相同的 DLCI 并不表示同一个虚连接。帧中继网络用户接口上最多可支持 1024 条虚电路,其中用户可用的 DLCI 范围是 16~1007,另外的 DLCI 代表特殊的功能,如 DLCI 0 和 1023 为 LMI 协议专用。

2) DLCI 地址影射

帧中继地址映射是把对端设备的协议地址(对端 IP 地址)与对端设备的帧中继地址(本地的 DLCI)关联起来,以便高层协议能通过对端设备的协议地址寻址到对端设备。帧中继主要用来承载 IP,在发送 IP 报文时,由于路由表只知道报文的下一跳地址,所以发送前必须由该地址确定它对应的 DLCI。这个过程可以通过查找帧中继地址映射表来完成,因为地址映射表中存放的是对端 IP 地址和下一跳的 DLCI 的映射关系。地址映射表可以通过手工配置 MAP,将 DLCI 号映射到远端的网络层地址。也可以由 Inverse ARP(逆向地址解析协议)动态维护。

在图 4.10 中,路由器管理者配置了一个 MAP,建立了 IP 地址为 172.16.10.3 和 DLCI 值为 48 的 PVC 的映射。

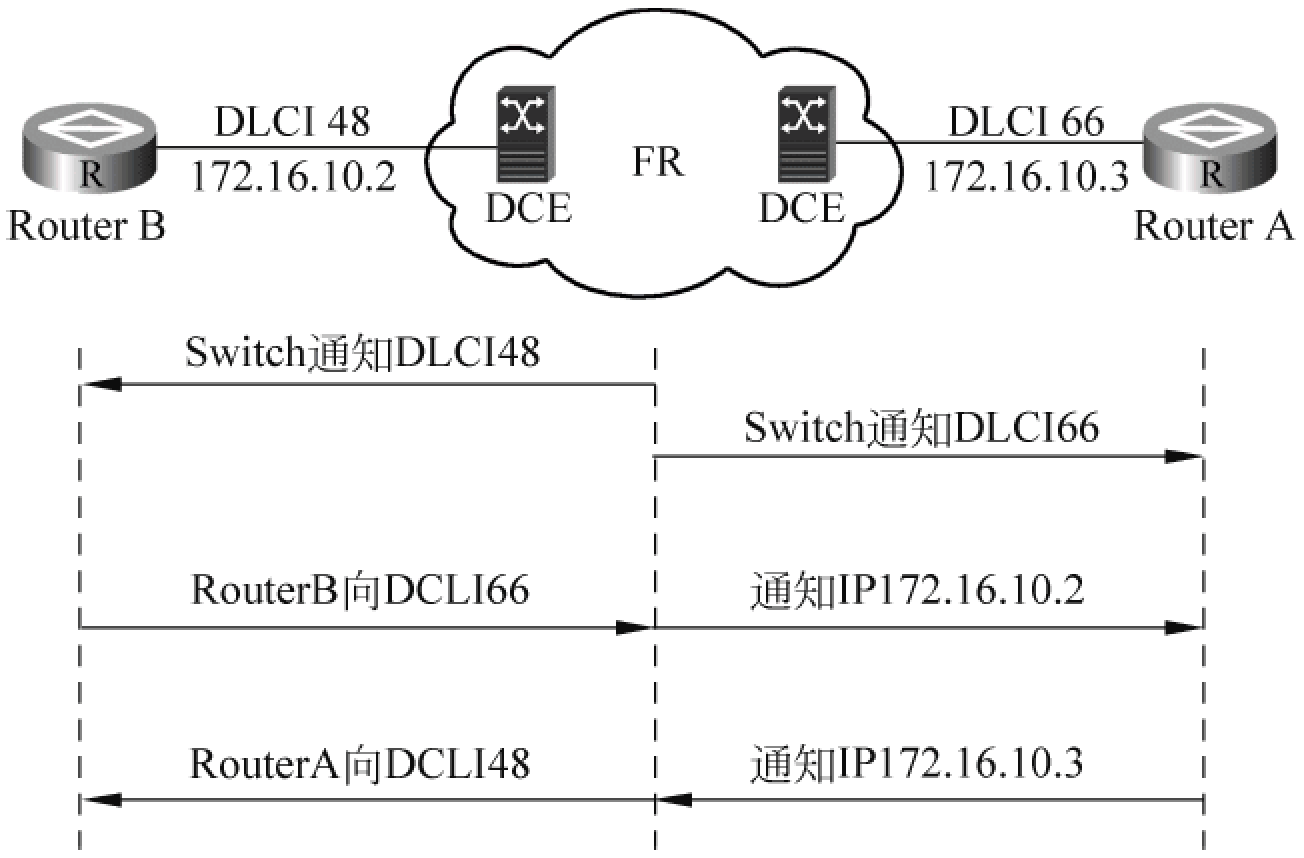


图 4.10 Inverse ARP 过程

也可以通过 Inverse ARP 进行地址映射。Inverse ARP 的主要功能是求解每条虚电路连接的对端设备的协议地址。如果知道了某条虚电路连接的对端设备的协议地址,在本地就可以生成对端协议地址与 DLCI 的映射(MAP),从而避免手工配置地址映射。

Inverse ARP 每当发现一个新的虚电路时(前提是本地接口上已配置了协议地址),就在该虚电路上发送 Inverse ARP 请求报文给对端,该请求报文包含有本地的协议地址,对端

设备收到该请求时，可以获得本地的协议地址，从而生成地址映射，并发送 Inverse ARP 响应报文进行响应，这样本地同样生成地址映射。

需要注意的是，如果已经手工配置了静态 MAP 或已经建立了动态 MAP，则无论该静态 MAP 中的对端地址正确与否，都不会在该虚电路上发送 Inverse ARP 请求报文给对端，只有在没有 MAP 的情况下才会向对端发送 Inverse ARP 请求报文。如果在 Inverse ARP 请求报文的接收端发现对端的协议地址与本地配置的 MAP 中的协议地址相同，则不会生成该动态 MAP。

4.3.2 FR 网络

1. 帧中继网络结构

网络服务商为用户提供固定的虚电路连接，用户可以申请许多虚电路，通过帧中继网络交换到不同的远端用户，现在比较常用的是帧中继的 PVC 业务。FR 广泛用在 WAN 中，能支持多种数据型业务，如 LAN 互连等，并可在分组交换数据网上提供高速传输业务。目前，最多的需求是帧中继支持图像传输和 LAN 互连。典型的帧中继网络结构如图 4.11 所示，它支持各类用户的接入，包括在用户侧的 T1/E1 复用设备、路由器、前端处理机、帧中继接入设备等，有时也称这些设备为室内用户设备(CPE)。网内设备专用于帧中继业务的帧中继节点。

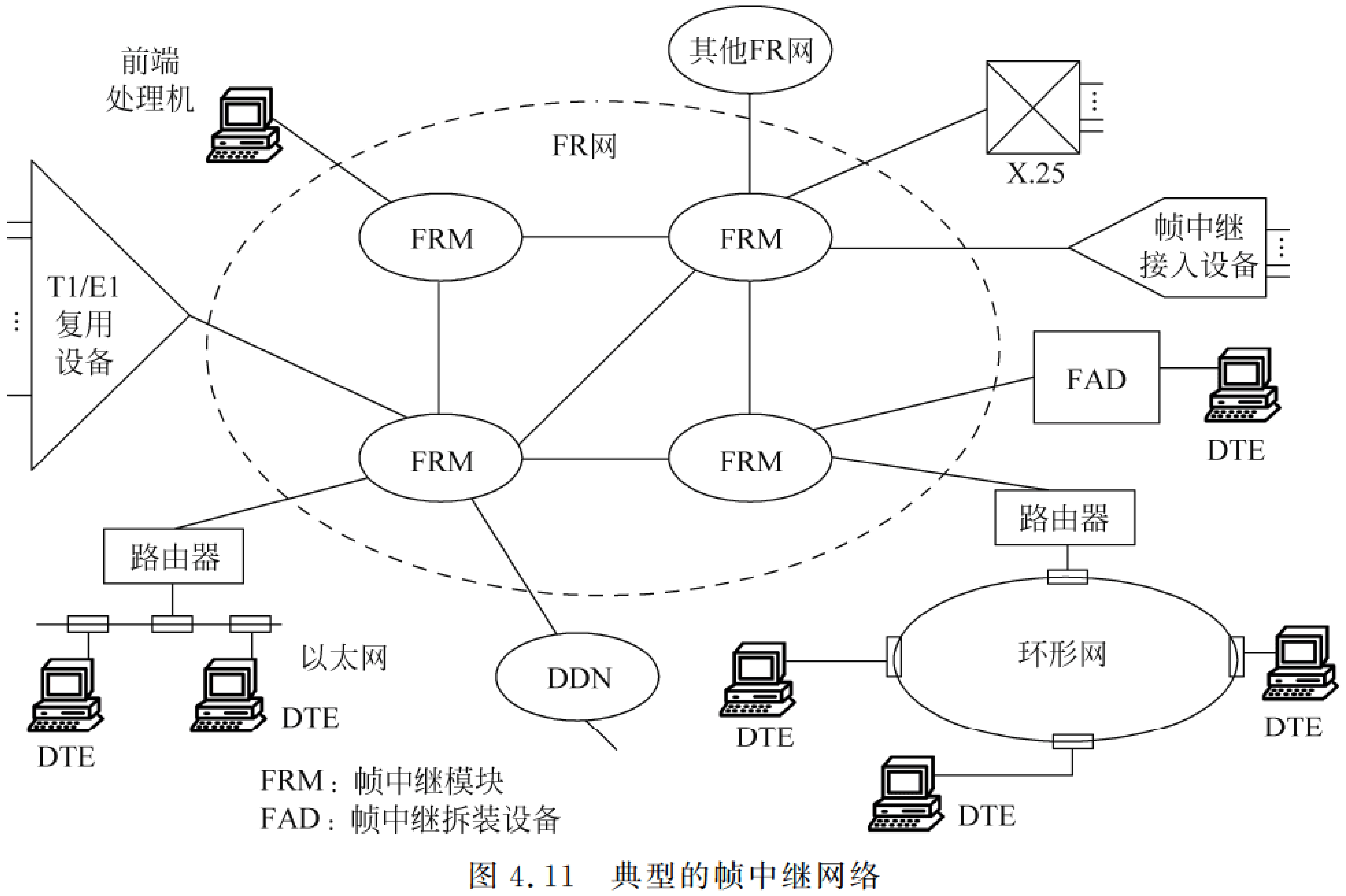


图 4.11 典型的帧中继网络

(1) 二线(或四线)话带调制解调传输方式：目前最高速率通常可达 38.4Kbps，有的高速调制解调器具有复用/分路功能，可为多个用户提供服务。这种方式主要适应要求速率低、距离较远的帧中继业务。

(2) 基带传输方式：二线或四线传输，传输速率通常为 16Kbps、32Kbps、64Kbps。具有时分复用的功能，将低于 64Kbps 的子速率复接到 64Kbps 的数据通路上，同时为多个用户提供接入服务。

(3) 2B+D 线路终端(LT)传输方式：采用 ISDN 数字用户环路技术，在一对双绞线上进行双向数据传输，可为多个用户提供接入，适合于较近的用户(6km 之内的用户)接入使用。

(4) ISDN 拨号接入方式：ISDN 拨号接入方式指 ISDN 用户终端通过拨号经 ISDN 接入到帧中继网络，一方面可以共享接口，降低成本；另一方面提高了接口的灵活性，扩展了 RF 的使用范围。

(5) PCM 数字线路传输方式：可利用接到用户处的光缆、微波数字电路等，并可以和其他业务合用，占用一条或多条 2M 位链路接入帧中继网。

(6) 其他方式：如 HDSL、ADSL 等接入到帧中继网络。

帧中继业务有两类：一类是具有 CCITT Q.922“帧方式承载业务 ISDN 数据链路层规范”接口的用户，称为帧中继用户；另一类是不具有 Q.922 接口的用户，称为非帧中继用户。帧中继用户可直接与 FRM 连接，非帧中继用户必须经帧中继拆装单元(FAD)及协议转换后才能与 FRM 相连。FRM 执行帧中继功能，即按照帧中继路由表和每个帧的帧头中数据链路连接标识符存储转发帧。由于 FRM 与 FAD 之间的专用电路可以独立于 DDN 节点和网络拓扑，所以可把帧中继业务看作是在专用电路上的增值业务和独立的帧中继网络(增值网络)。

帧中继业务主要作为一种承载业务应用在 WAN 中，支持多种数据型用户业务。可以通过在 DDN 节点内引入帧中继模块(FRM)来提供帧中继业务。根据帧中继业务的需求，可选择在某些 DDN 节点内设置 FRM，利用专线电路连通 FRM，所配置的专线电路专供帧中继业务使用。用户以一条专线接入 DDN 可以同时与多个点建立帧中继电路(PVC)。所以，在这种情况下，帧中继业务是由建立在 DDN 之上，逻辑上又独立于 DDN 的帧中继业务网络来提供的，可以认为 DDN 上存在一个虚拟的帧中继网络。这时帧中继用户的入网速率为(9.6～2048)Kbps，即 9.6Kbps，14.4Kbps，16Kbps，19.2Kbps，32Kbps，48Kbps，$N\times$64Kbps(N=1～32 可选)。

帧中继网主要满足传输速率为 64Kbps～45Mbps 的通信需求。它将作为 CHINAPAC、INANET 等的中继汇接、高速数据用户互连，满足远程教学、医疗、设计、服务及多媒体等宽带业务的需求。用户设备和网络接口设备之间的物理接口，通常提供以下之一的接口规程。

(1) X 系列接口，如 X.21 接口、X.21bis 接口等。

(2) V 系列接口，如 V.35、V.36、V.10、V.11、V.24 等。

(3) G 系列接口，如 G.703，传输速率可为 2Mbps、8Mbps、34Mbps 或 155Mbps 等。

(4) I 系列接口，如支持 ISDN 传输基本速率接入的 I.430 接口和支持 ISDN 一次群传输速率接入的 I.430 接口等。

2. 帧中继组网技术

针对帧中继技术，简单地归结为以下几点。

(1) 帧中继简化了分组交换协议，使节点处理负担大大减轻，使帧中继网的传输速率大幅度提高，速率最高可达 E3(34Mbps)，而且节点延时降低到了 2ms 以下。

(2) 帧中继节点不执行差错控制，并简化了节点机之间的确认方式。节点不需要保存待重发的信息，信息帧在网中直接通过，从而进一步提高了传输速率，并减少了网络时延。

(3) 帧中继使用统计时分复用，在信道上动态划分数据链路，达到带宽动态适配的需

求。这种技术向用户提供共享的网络资源,每一条线路和网络端口都可由多个终端按信息流量共享,从而大大提高了网络资源利用率。

(4) 帧中继简化了节点机间的协议处理,将更多的带宽留给用户数据,因而能提高吞吐量,降低时延。

(5) 帧中继对于经常需要传送突发数据业务的用户来说十分有效。例如,当某一用户的业务突发时,其他用户可能没有或只有少量数据,则该用户就可以多占 CIR(承诺的信息速率)之外的带宽,而只需付预定带宽的费用。因此,帧中继业务可以共享网络资源,较为经济。

最后要给出的是一个分组交换、帧中继和 DDN 的性能比较表,见表 4.1,供读者学习参考。

表 4.1 数据通信网络性能比对表

	X.25	FR	DDN
OSI	1～3 三层	下二层	物理层
复用方法	动态复用	动态复用	静态复用
所用协议	X.25 等	Q.933 等	无
差错控制	检查、重发	只检查	无
虚电路	SVC、PVC	PVC、SVC(无)	无(TDM)
DTE 速率	64Kbps 2.4Kbps 9.6Kbps 4.8Kbps	2Mbps $N\times$64Kbps 9.6Kbps (8～10)Mbps 等	2Mbps $N\times$64Kbps 9.6Kbps
中继最大速率	56/64Kbps	(2～34)Mbps	2Mbps
典型应用场合	交互式短报文	局域网互联	专线用户
网络间标准	X.75	NNT	无
流量控制	第二/第三	高层协议	无

4.4 异步传输模式(ATM)

ITU-T 在 I.113 建议中定义:ATM 是一种传递模式,在这一模式中,信息被组织成信元(cell),包含一段信息的信元不需要周期性地出现,从这个意义上讲,此传递模式是异步的。

4.4.1 ATM 交换

1. ATM 技术的特点

ATM 作为 ITU-T 建议的 B-ISDN 的传递方式,具有以下技术特点。

(1) ATM 是一种统计时分复用技术,它将一条物理信道划分为多个具有不同传输特性的虚电路提供给用户,实现网络资源的按需分配。

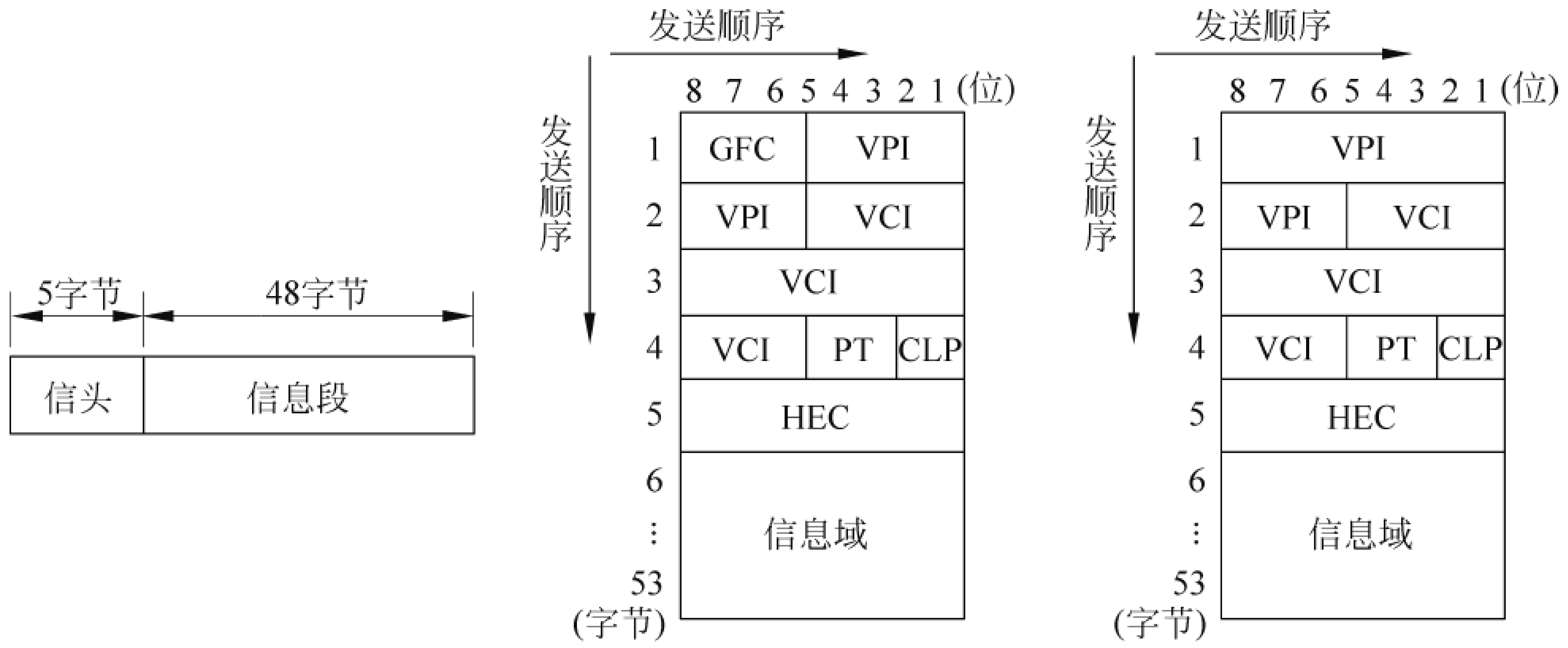
(2) ATM 利用硬件实现固定长度分组的快速交换,具有时延小、实时性好的特点,能够满足多媒体数据传输的要求。

(3) ATM 是支持多种业务的传递平台，并提供服务质量保证，ATM 通过定义不同的 AAL(ATM 适配层)来满足不同业务对传输性能的要求。

(4) ATM 是面向连接的传输技术，在传输用户数据之前必须建立端到端的虚连接，所有数据，包括用户数据，信令和网管数据都通过虚连接进行传输。永久虚连接(PVC)可以通过网管功能建立，但交换虚连接(SVC)必须通过信令过程建立。

2. ATM 信元结构

ITU-T 在 I. 361 建议中有规定，一个 ATM 信元长 53 字节，前面 5 字节称为信头(header)，后面 48 字节称作信息段(payload)，是用户数据。ATM 信元实际上是将语音、数据及图像等所有的用户数字信息分解成固定长度的数据块，并保存数据块前装配地址、丢失优先级流量控制、差错控制(HEC)信息等，形成了 5 字节的信元头。这样再加上 48 字节的用户数据信息就构成一个了完整的 ATM 信元，ATM 的信元结构如图 4. 12(a)所示。ATM 的 5 个字节信头中包含流量控制信息、虚连接标识符、净荷类型、信元丢失优先级和信头差错控制信息。它们对应于 ATM 两种不同信头的信元格式，如图 4.12(b)、(c)所示。



(a) 信元结构　(b) 用户-网络接口　(c) 网络节点接口

图 4.12　ATM 信元结构和信头格式

(1) 一般流量控制(GFC)：由 4 位组成，仅用于 UNI。不用时可置为 0000，可以用于以后的流量控制，或在共享媒体的网络中表示不同的接入。

(2) 虚通路标识符(VPI)：用户-网络接口由 8 位组成，用于路由选择，可标识 256 个 VP；而在网络节点接口由 12 位组成，以增强网络中的路由选择功能，可标识 4096 个 VP。

(3) 虚信道识别符(VCI)：由 16 位组成，可标识 65 536 个 VC，用于 ATM 虚信道路由选择。VPI/VCI 一起标识一个虚连接。

(4) 消息类型(PT)：该字段的长度为 3 位。用于标识净荷的类型，第三位“0”表示为数据信元，为“1”表示为维护操作(OAM)信元；对数据信元，第二位用于前向拥塞指示，第一位用于 AAL5(ATM 适配层)；对 OAM 信元，后两比特表明了 OAM 信元的类型。

(5) 信元丢失优先级(CLP)：该字段由 1 位组成，用于表示信元丢失的等级，用于拥塞控制。CLP=0，网络尽力为其提供带宽资源，以防信元丢失；CPL=1，可根据带宽情况丢弃信元。

(6) 信头差错控制(HEC)：该字段是长度为 8 位的 CRC 校验码，可提高信头的传输可靠度，用于检测信头的比特差错和信元定界。根据 ITU-T 的 I. 432 建议，这一区域的处理

在物理层进行。

3. ATM 参考模型结构

图 4.13 所示为 ATM 参考模型与 OSI 协议的对应关系，主要分为物理层(PHY)、ATM 层、AAL 层和高层，以下将对各层展开介绍。ATM 参考模型有时也称 B-ISDN 参考模型。

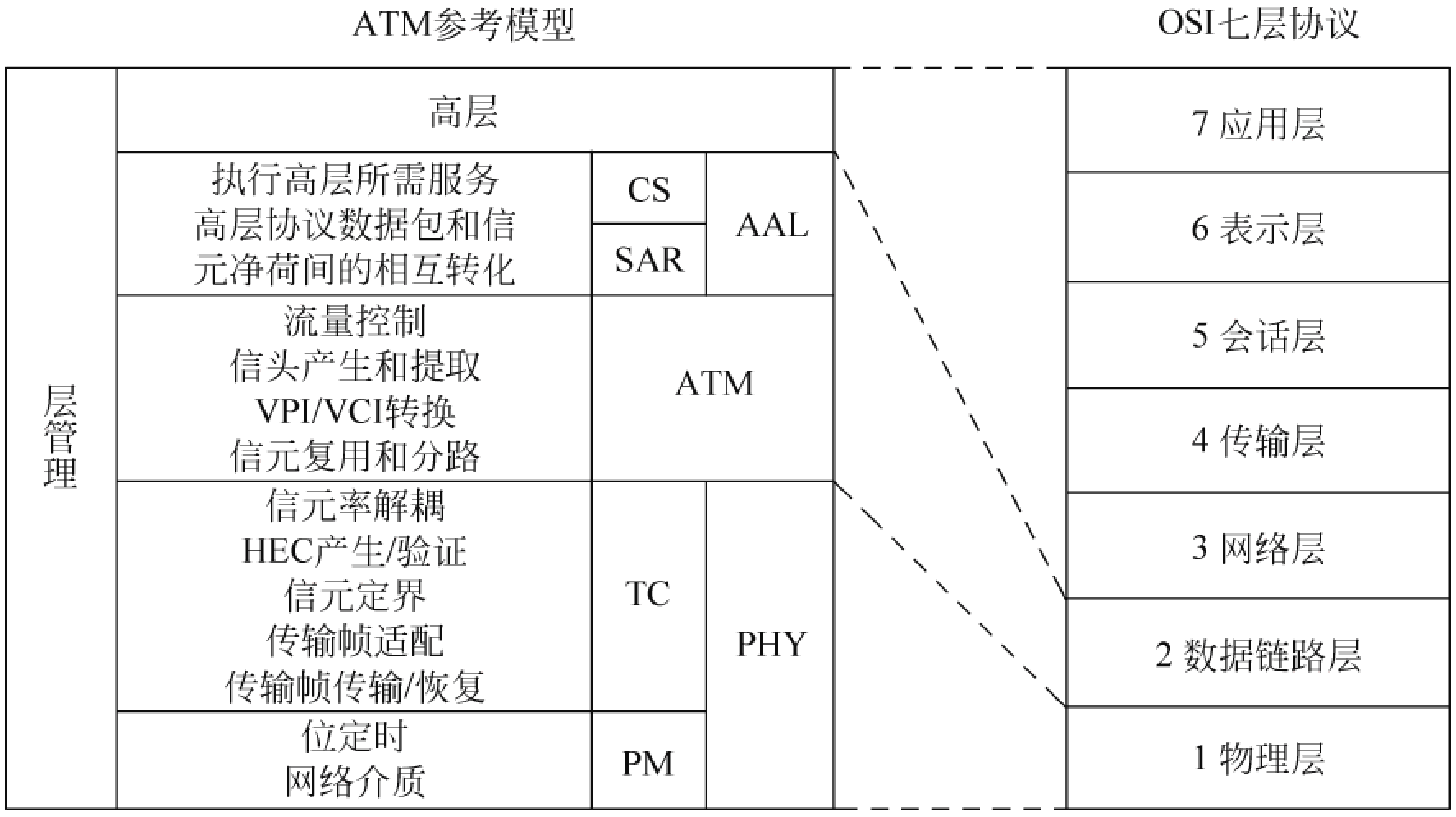


图 4.13 ATM 参考模型和 OSI 协议对应关系

1) 物理层

物理层完成传输信息(位/信元)功能，负责制定物理媒体的工作，利用通信线路的比特流传送功能实现传送 ATM 信元的功能，它包含两个子层：传输汇聚子层(Transmission Convergence sublayer，TC)、物理媒介子层(Physical Medium sublayer，PM)，这两个子层交换的数据要互相同步。物理层主要处理和完成相邻 ATM 层间 ATM 信元的传输，向 ATM 层提供传输接入。在传输方向上，ATM 层把 ATM 信元(HEC 值除外)传递到物理层。在接收方向上，ATM 从物理层接收 53 字节的信元。

2) ATM 层

ATM 层负责交换、路由选择和信元复用。ATM 层为 ATM 适配层和物理层之间提供接口，ATM 层只涉及信元的信头功能，而不处理信息域的信息类型、业务时钟频率信息。以下介绍 ATM 层的主要功能。

(1) 信元的复用和解复用：信元复用/解复用在 ATM 层和物理层的 TC 层接口处完成，在源点负责对多个虚连接的信元进行复接和在目的端对接收的信元进行分解。

(2) 有关信头的操作：信头操作在用户端为填写 VPI/VCI 和 PT 负责源点产生信头和宿点翻译信头，在网络节点中为 VPI/VCI 翻译，负责在每个 ATM 节点对信头进行标记/识别。

(3) 一般流量控制(GFC)：GFC 是指信头的 GFC 控制位，用于控制终端到网络的业务流量。

(4) ATM 的交换：信元进入交换节点后，信头中的 VPI/VCI 被迅速提取并读出，然后查找翻译表(路由连接表和标识转换表)，将它们的新值填入，信元被送入新值所对应的 VC/VP 链路输出完成交换。

3）AAL 层

AAL 层完成将各种业务的信息适配成 ATM 信元流。目前，ITU-T 已提出 4 种不同的 AAL 层协议，以支持 ATM 网的 4 类业务，记为 AAL1、AAL2、AAL3/4、AAL5，业务上分别称为 A、B、C、D 类。AAL 业务分类见表 4.2，另外还有 X、Y 两类业务。以下简单介绍 A、B、C、D 类业务。

表 4.2　AAL 业务分类

业务分类	A	B	C	D
AAL 分类	AAL1	AAL2	AAL3/4	AAL5
端到端定时	要求		不要求	
比特率	固定		可变	
连接方式	面向连接		无连接	
业务举例	固定比特率的语音、动态图像等	可变比特率的动态图像	数据传输	通过 WAN 的两 LAN 间数据传输

（1）A 类。固定的比特率（CBR）业务，是对应源点和终点间具有定时关系的固定比特率面向连接业务，两路典型的 A 类业务是 64Kbps 语音和 CBR 视像。

（2）B 类。可变比特率（VBR）业务，支持面向接续业务，支持延时敏感性的业务，如移动电话业务（13Kbps 或 9.5Kbps）、短分组或低速数据等。

例：若要求传送 13Kbps 的移动电话业务，形成一个信元所需要的时间为：48（字节）×8（位）/13（Kbps）＝29.5ms。也就是说，在这个间隔时间内可以复用 155.520（Mbps）×29.5(ms)/48(字节)/8(位)＝11 963 个其他信元。在这里，ATM 网的传输速率为 155.520Mbps。

（3）C 类。面向接续数据业务，其接续是在数据被传输以前建立的，其速率是可变的，但没有介质延迟。AAL3/4 规程用于 C 类和 D 类业务，如 X.25、帧中继、局域网等。

（4）D 类。无接续数据业务，此业务是对应源点和终点间不具有定时关系的可变比特率无接续业务，通过 WAN 的两个局域网间无接续数据传递是此类业务的典型实例。

4）高层

根据不同的业务特点完成高层功能。

4．VC/VP 交换原理

所谓交换，是指输入的 ATM 逻辑信道到输出 ATM 逻辑信道的信息交换，ATM 逻辑信道有两个特征：一个是以物理端口号表征的物理出线/入线；另一个是以 VCI、VPI 表征的建立在物理端口之上的逻辑信道。所谓复用（集中），就是将不同的 ATM 的虚通路统计合并成单一的 ATM 信息流，“集中”更强调将入线的数量集中到更少量的入线上。而扩展（分路）是复用（集中）的逆操作。

例如，如图 4.14 所示，用户 A 要向用户 B 发送数据，那么这个数据在由用户 A 发送到它的 ATM 终端设备时，在发送端的 ATM 终端设备完成了用户数据到信元的转化，然后再发送到 ATM 网络设备进行 ATM 交换。ATM 网络设备根据信元的 VCI 和 VPI 值，查询端口路由表，计算出在接收端的 ATM 网络设备中与自己的端口相对应的 VCI 和 VPI 值，并将信元 VCI 和 VPI 的值变得与其一致，然后发送到出端口。接收端的 ATM 网络设备根据同样的原理找到接收端的 ATM 终端设备，改变信元 VCI 和 VPI 的值将其发送。接收端

的 ATM 终端设备完成信元到用户数据的转化，交给用户 B。但是在交换之前，两端必须建立一个连接。ATM 的交换本质是分组交换。ATM 交换分为 VP 交换和 VC 交换两种。VPI 和 VCI 仅在两个物理节点间具有局部意义。一般情况下，VP 交换只改变 VPI 的值不改变 VCI 的值；VC 交换既改变 VPI 的值又改变 VCI 的值。

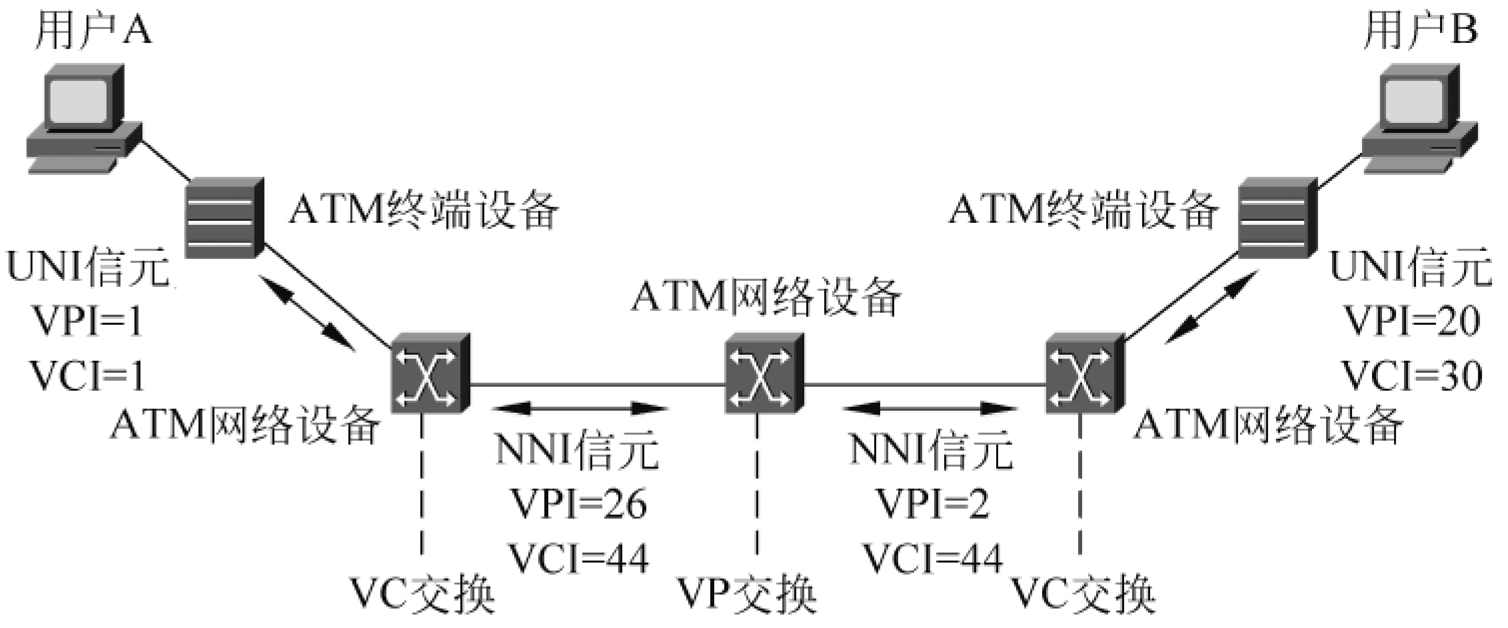


图 4.14 VC/VP 交换原理

A 发出的信元在要经过的 ATM 节点上建立一系列的交换表，保证信元经过逐次转发后最后能到达 B。这些交换表建立后，就形成了信元 A 到 B 所经过的路径，实际上就是我们前面所说的 ATM 虚连接(或虚信道)，根据建立方式的不同可以分为交换虚连接(SVC)和永久虚连接(PVC)。

4.4.2 ATM 网络

1. ATM 网络及接口

在网络结构上，ATM 网可分为 3 种：公用 ATM 网、专用 ATM 网和接入 ATM 网。ATM 网之间及与终端设备之间通过各种接口进行互连，如图 4.15 所示，其主要接口功能介绍如下。

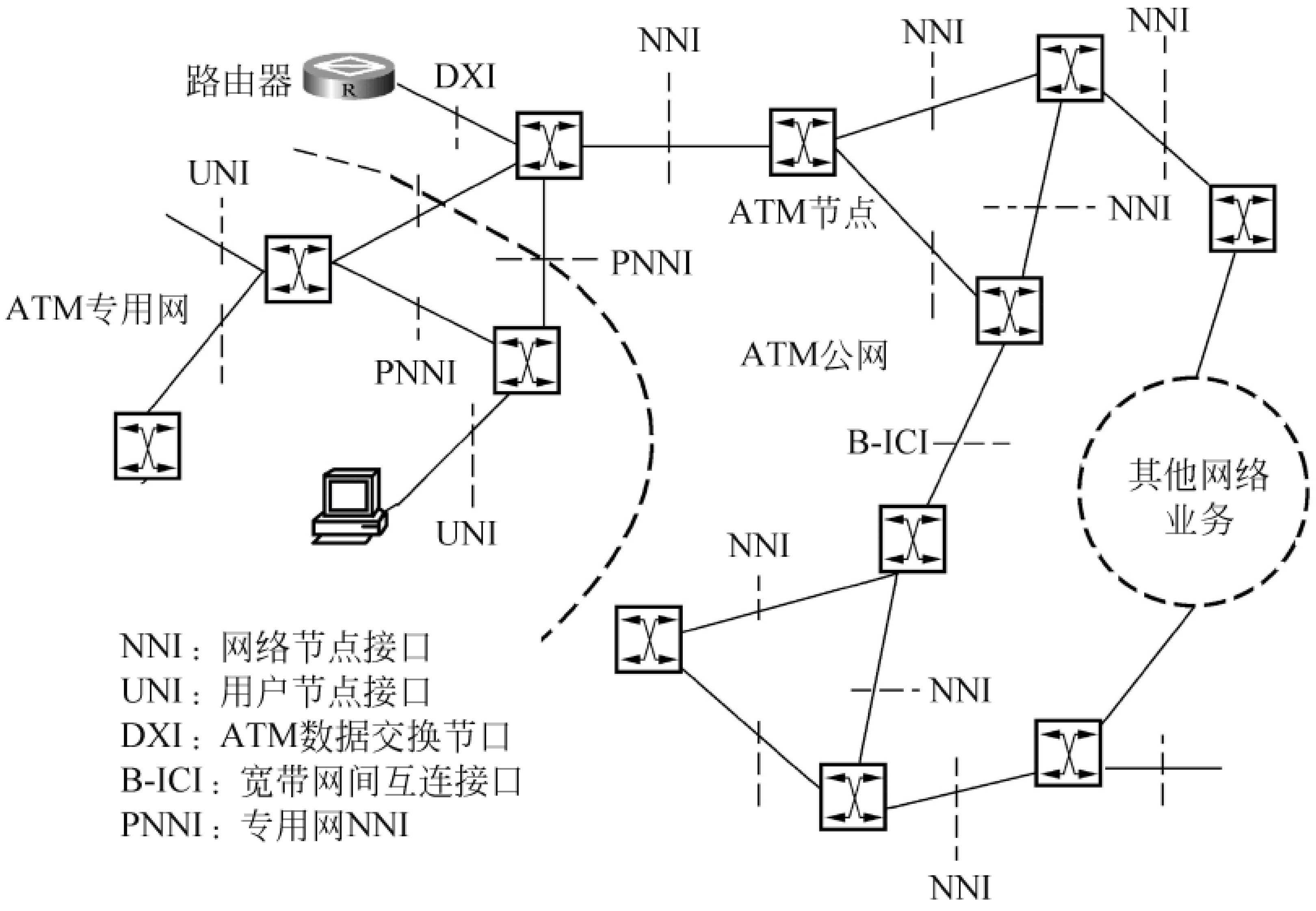


图 4.15 ATM 网络及接口

1）UNI

UNI(User-Network Interface)为 ATM 网中的用户-网络接口，它是用户设备与网络之间的接口，直接面向用户。UNI 定义了物理传输线路的接口标准，即用户可以通过怎样的物理线路和接口与 ATM 网相连，还定义了 ATM 层标准、UNI 信令、OAM 功能和管理功能等。

2）NNI

NNI(Network to Network/Network Node Interface)可理解为网络节点接口或网络-网络之间的接口，一般为两个交换机之间的接口，与 UNI 基本一样，但由于 NNI 关系到连接在网络的路由选择问题，所以特别对路由选择方法做了说明。

3）B-ICI

B-ICI(B-ISDN Inter-Carrier Interface)为宽带网间接口，它的定义基于 NNI，侧重于不同经营者的 ATM 间的接口。其特点是支持不同网络间的多种业务传送。

4）DXI

数据交换接口(Data Exchange Interface，DXI)定义在数字终端设备(DTE)和数字连接设备(DCE)之间，DTE 通过 DXI 与 DCE 相连，再通过 ATM UNI 接入 ATM 网中。

5）FUNI

FUNI(Frame-based UNI Interface)将 ATM 适配功能完全移入了交换机内部，终端和 ATM 交换机之间传送 FUNI 帧，所以与基于信元的 DXI 接口相比，它在接入线上有更高的效率。

2．IP Over ATM(IPOA)

IPOA 网络结构如图 4.16 所示，在这种情况下，虚拟局域网的所有终端都连接到 ATM 网络上，任意两个有通信要求的 IPOA 终端间必须建立 PVC，ATM 网络的网关是一个路由器。网络内的地址解析都通过连接到内网的 ARP 服务器来实现。

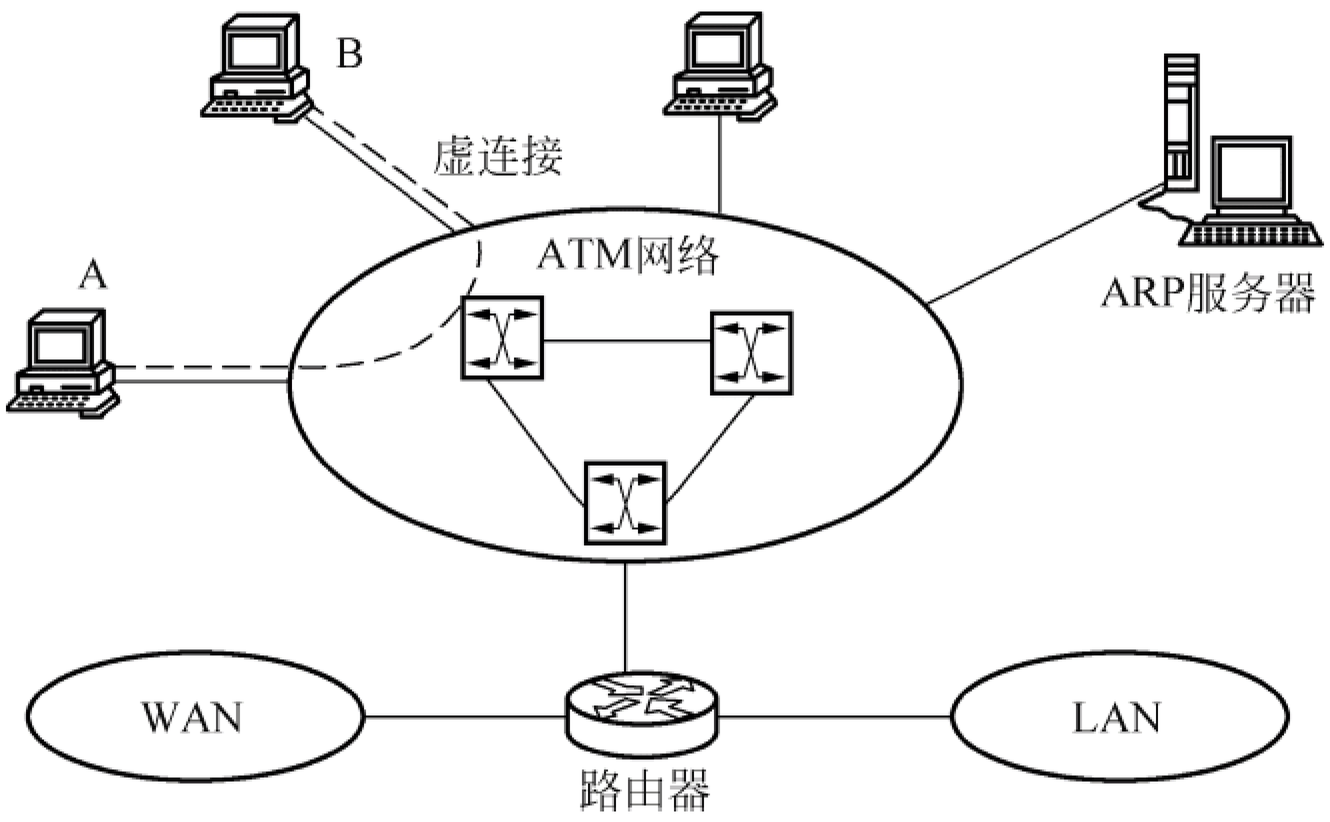


图 4.16 IPOA 网络结构

主机 A 和主机 B 间如果要实现 IP 数据交互，首先要配置一条 PVC。主机 A、B 分别将自己的 IP 地址回应给对方，于是双方均在自己的系统建立对方的 IP 地址和 PVC 的映射表，当然，该表中还有其他作出回应的主机 IP 地址和 PVC 的映射。主机 A 如果有目的地为主机 B 的 IP 数据包，主机 A 的 ATM 层在收到该数据包转化后的 AAL-PDU 就会根据

主机 B 的 IP 查找 IP 地址和 PVC 的映射表，索引出相应的 PVC，并据此填写 ATM 信元头中的 VPL/VCI。该信元即可由 ATM 网络传送到主机 B。ATM 交换是信元交换，它结合了电路交换和分组交换的优点，是面向连接的。SVC 通过信令建立和拆除连接，PVC 相当于半永久连接，VPI/VCI 表由人工维护。

最后，ATM 的基本优点可以概括为：能够根据用户需求分配带宽；网络可扩展性好，ATM 不仅用于公用网，而且在专用网、LAN 中都有很好的、广泛的应用；信息段被透明传送，免除了差错控制和流量控制，简化了网络控制；面向连接的工作方式；高信道使用率；ATM 可以提供多业务，只用 ATM WAN 就可以提供本网 ATM 业务、帧中继(FR)业务、交换型多兆位数据业务(SNDS)；ATM 支持 QoS 的技术，ATM 业务分类使用户有一系列可选项目，在费用、性能之间做出折中，可保证现行网络的应用和未来通信业务网络的发展。

习题

一、单选题

1. 被称作分组数据网的枢纽的设备为(　　)。

A. 分组终端　　B. 分组拆装设备

C. 分组交换机　　D. 网络管理中心

2. 帧中继技术主要用于传递(　　)业务。

A. 数据　　B. 语音　　C. 视频　　D. 综合多媒体

3. 信元是一种固定长度的数据分组。一个 ATM 信元长(　　)。

A. 53 字节，前 5 字节称为信头，后面 48 字节称为信息域

B. 64 字节，前 8 字节称为信头，中间 48 字节称为信息域，后 8 字节称为信尾

C. 48 字节，前 5 字节称为信头，后 43 字节称为信息域

二、多选题

1. X.25 网络包含了三层，分别为(　　)，和 OSI 参考模型的下三层一一对应，它们的功能也是一致的。

A. 传输层　　B. 物理层

C. 数据链路层　　D. 网络层(分组层)

2. DDN 由(　　)组成。

A. 用户环路　　B. DDN 节点

C. 数字信道　　D. 网络控制管理中心

3. ATM 参考模型中，ATM 层的主要功能有(　　)。

A. 信元复用/解复用　　B. 信元 VPI/VCI 翻译

C. 信元头的产生和提取　　D. 一般流量控制功能

三、是非判断题(将正确的题打√)

1. 面向连接的服务，具有连接建立、数据传输、连接释放 3 个阶段。

2. DDN 所提供的数据信道是半永久的，是交换型的。

3. 帧中继技术是在 OSI 的第二层上用简化的方法实现传送和交换数据单元的技术。

4. 异步转移模式(ATM)是一种基于分组交换和时分复用的技术。

四、简答题

1. 简述 DDN 的组成。

2. 什么是分组交换虚电路?

3. X.25 协议使用虚电路号被划分成哪 4 个区域?

4. 如何理解 DLCI? 它在帧中继中起到什么作用?

5. 为什么说“ATM 技术融合了电路交换传送模式,采用了时分统计复用和信息分组的设计思想发展而成”?

6. 说明 ATM 的 VC/VP 交换原理。

7. 简述 DDN、分组交换、帧中继、ATM 的技术特点。

8. 详述 ATM 网络中的主要接口名称,以及各个接口在 ATM 网络中的作用。

第5章 路 由 器

路由器主要功能是实现路由选择与网络互连，即通过一定途径获得网络拓扑特性，并通过一定的路由算法获取到达各网络的最佳路径，建立相应路由表，从而将每个IP包经过跳到跳(hop to hop)传到目的地。本章主要介绍路由器的工作原理及其有关配置。

5.1 路由器技术

所谓路由，就是指导IP数据包发送的路径信息。路由器是一种用于网络互连的计算机设备，它工作在OSI参考模型的网络层，为不同的网络之间报文寻找路径并存储转发，它提供了网络互连的机制，实现将数据包从一个网络发送到另一个网络。

5.1.1 路由器基本概念

1. 路由分类

在互联网中，路由器根据所收到的数据报头的目的地址，选择一个合适的路径，将数据包传送到下一个路由器。数据包在网络上的传输过程，就好像是在体育运动中的接力赛一样，每一个路由器只负责自己本站数据包通过最优的路径转发，随后通过多个路由器的接力，将数据包转发到目的地，路径上最后的路由器负责将数据包送交到目的子网或主机。根据路由的目的地不同，可以划分为子网路由和主机路由；根据目的地与该路由器是否直接相连，又可分为直接路由和间接路由。因此，路由可以概括为以下四类。

(1) 子网路由：目的地为子网。

(2) 主机路由：目的地为主机。

(3) 直接路由：目的地所在网络与路由器直接相连。

(4) 间接路由：目的地所在网络与路由器不是直接相连。

一般路由器都会运行一些动态路由协议，以实现动态寻径。有些路由器还会支持两种以上的网络协议，以支持异种网络互连，具有网关的功能。路由器必须具备两个或两个以上的接口，用于连接不同的网络(在现实网络中也存在只有一个接口的情况，这种方式的路由器称为独臂路由器，但应用不多)。路由器具有存储、转发、寻径功能，实现速率匹配与路由寻径。

2. 路由器的作用

路由器的主要作用是将不同的网络互连为一个整体，它的主要作用表现在以下几个方面。

(1) 数据转发：路由器必须具有根据数据分组的目的网络地址转发分组的功能。

(2) 路由(寻径)：为了实现数据转发，路由器必须有能力建立、刷新路由表，并根据路

由表转发数据包。

(3) 备份、流量流控：为了保证网络可靠运行，路由器一般都具备主备线路的切换及流量控制功能。

(4) 速率适配：不同接口具有不同的传输速率，路由器可以利用自己的缓存及流控协议进行适配。

(5) 隔离网络：路由器可以隔离和防止广播网络，同时也可以对数据包施行灵活多样的过滤策略，以保证网络安全(防火墙作用)。

(6) 异种网络互连：互联网的初衷就是为了实现异种网络互连，现代路由器一般都会实现两种以上的网络协议以实现异种网络互连(相当于网关作用)。

3. **路由表**

路由器依据路由表来为报文寻径，路由表(routing tables)由路由协议建立和维护，路由协议的设计则是依据某种路由算法。路由表信息主要内容项如图 5.1 所示，标准的路由表最初是一个二维组(目的网络地址、下一跳地址)，其中不携带子网信息，不能满足子网寻径。引入子网编址以后，路由表加入子网掩码，于是路由表变为三维组：子网掩码、目的网络地址和下一站地址。

Destination/Mask	proto	pref	Metric	Nexthop	Interface
0.0.0.0/0	Static	60	0	120.0.0.2	Serial0
8.0.0.0/8	RIP	100	3	120.0.0.2	Serial0
9.0.0.0/8	OSPF	10	50	20.0.0.2	Ethernet0
11.0.0.0/8	Static	60	0	120.0.0.2	Serial0
20.0.0.0/8	Direct	0	0	20.0.0.1	Ethernet0
20.0.0.1/32	Direct	0	0	127.0.0.1	LoopBack0

图 5.1　路由表显示信息

路由器转发数据包的关键是路由表。每个路由器中都保存着一张路由表，表中每条路由项都指明数据包到某子网或某主机应通过路由器的哪个物理接口发送，然后就可到达该路径的下一个路由器，或者不再经过别的路由器而传送到直接相连的网络中的目的主机。

路由表主要包含下列关键项。

目的地址(Destination)：用来标识 IP 包的目的地址或目的网络。

网络掩码(Mask)：与目的地址一起来标识目的主机或路由器所在的网段的地址。将目的地址和网络掩码“逻辑与”后可得到目的主机或路由器所在网段的地址。

输出接口(Interface)：说明 IP 包将从该路由器的哪个接口转发。

下一跳 IP 地址(Nexthop)：说明 IP 包所经过的下一个路由器接口应用的 IP 地址。

在路由表中有一个 Protocol 字段，指明了路由的来源，即路由是如何生成的。路由的来源主要有以下 3 种。

(1) 链路层协议发现的路由(Direct)。开销小，配置简单，无需人工维护，只能发现本接口所属网段拓扑的路由。

(2) 手工配置的静态路由(Static)。静态路由是一种特殊的路由，它由管理员手工配置而成。通过静态路由的配置可建立一个互通的网络，但这种配置问题在于：当一个网络故障发生后，静态路由不会自动修正，必须有管理员的介入。静态路由无开销，配置简单，适合简单拓扑结构的网络。

(3) 动态路由协议发现的路由，如通过路由协议 RIP、OSPF 等得到的路由信息。当网络拓扑结构十分复杂时，手工配置静态路由工作量大而且容易出现错误，这时就可用动态路由协议，让其自动发现和修改路由，无需人工维护，但动态路由协议开销大，配置复杂。

在每一个协议栈中，都制定了一些路由协议创建路由表。例如，OSI 参考模型的 IS-IS (中间系统到中间系统的路由协议)；TCP/IP 协议栈的 RIP(路由信息协议)、OSPF(开放式最短路径优先)协议；IPX/SPX(分组交换/顺序分组交换)协议栈的 IPX 协议等。路由器能支持多个相互独立的路由协议，能为不同的网络协议栈相对应的路由协议维护各自的路由表。

4. 路由优先级

要到达相同的目的地，不同的路由协议(包括静态路由)可能会发现不同的路由，但并非这些路由都是最优的。事实上，在某一时刻，到达某一目的地的当前路由仅能由唯一的路由协议来决定。这样，各路由协议(包括静态路由)都被赋予了一个优先级(preference)，当存在多个路由信息源时，具有较高优先级(数值越小，表明优先级越高)的路由协议发现的路由将成为最优路由，并被加入路由表中。

不同厂家的路由器对于各种路由协议优先级的规定各不相同。如华为(Quidway)路由器的默认优先级见表 5.1。其中：0 表示直接连接的路由，255 表示任何来自不可信源端的路由。除了直接路由(Direct)外，各动态路由协议的优先级都可根据用户需求，手工进行配置。另外，每条静态路由的优先级都可以不相同。

表 5.1 路由器的默认优先级

路由协议或路由种类	相应路由的优先级
Direct	0
OSPF	10
Static	60
RIP	100
IBGP	130
OSPF ASE	150
EBGP	170
Unknown	255

5. 路由花费

路由花费(metric)表示到达这条路由所指的目的地址的代价，通常路由的花费值会受到线路延迟、带宽、线路占有率、线路可信度、跳数、最大传输单元等因素的影响，不同的动态路由协议会选择其中的一种或几种因素来计算花费值(如 RIP 用跳数来计算花费值)。该花费值只在同一种路由协议内有比较意义，不同的路由协议之间的路由花费值没有可比性，也不存在换算关系。静态路由的花费值为 0。

5.1.2 路由器构成

IP 包途经每个路由器时，需要排队、协议处理和寻址选择路由等各个软件处理环节，会造成一定的延时，这与它的构成是有关系的。

1. 路由器结构功能

路由器逻辑上由输入/输出接口、数据转发部分、路由管理部分、用户配置接口共 5 部分构成。路由器结构框图如图 5.2 所示，路由选择处理机就是通过路由表以及具体的交换结构完成数据转发部分等功能。

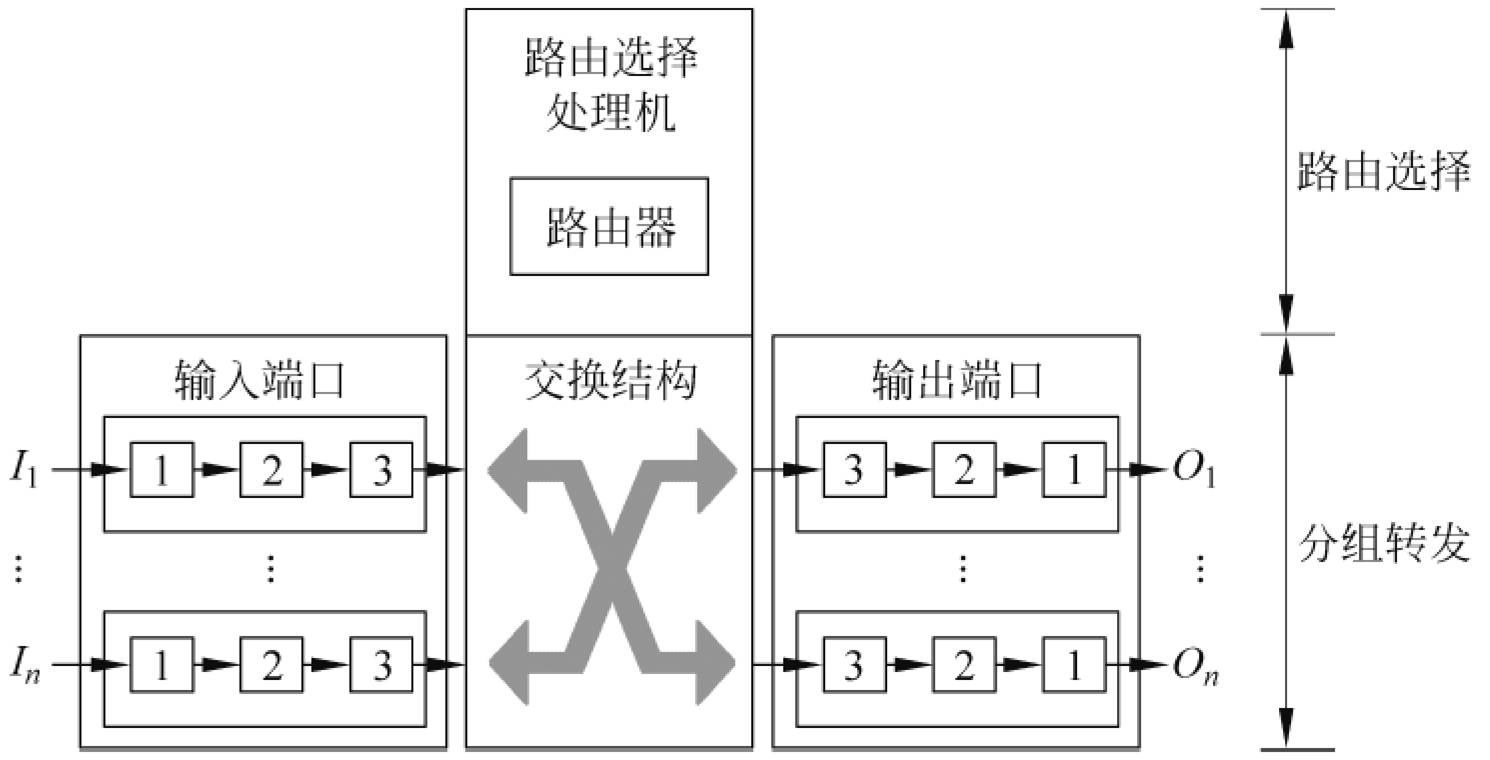


图 5.2 路由器结构框图

1—物理层；2—数据链路层；3—网络层

输入接口对分组的处理：物理层按位进行数据的接收和发送。数据链路层按链路层协议接收帧，并去掉帧的首部和尾部，将其送往网络层的队列中排队等待处理。排队等待会产生一定的时延。

输出接口对分组的处理：交换结构传送过来的分组先进行缓存。数据链路层处理模块将分组加上链路层的首部和尾部，交给物理层后送往线路。

物理层按位进行数据的接收和发送。数据链路层按链路层协议接收帧，并去掉帧的首部和尾部，将其送往网络层的队列中排队等待处理。排队等待会产生一定的时延。

单一网络的分组交换是基于查表机制。所查的表有路由表和转发表。

路由表(routing table)是根据某种路由选择算法得出的一张表格，供路由选择之用。路由选择协议负责搜索分组从某个节点到目的节点的最佳传输路由，以便构造路由表。

转发表(forwarding table)是根据路由表构造的一张表格。当交换节点收到分组后，根据其目的地址查找转发表，并找出应从节点的哪一个接口将该分组发送出去。平时并不严格区分转发表和路由表，在转发分组时既可以说“查找转发表”，也可以说“查找路由表”。

路由器或交换机属于分组交换设备，其结构类型如图 5.3 所示，分为基于存储器转发交换的存储器型、基于统计时分复用交换的总线型和基于空分交换的互联网络型。

图 5.3(a)是共享存储器型结构，这种结构依赖中心交换引擎来提供全端口的高性能连接，由核心引擎检查每个输入数据包以决定路由。这种方法需要较大的内存带宽、较高的管理开销，尤其是随着交换机端口的增加，内存的性价比成为实现核心交换的瓶颈。

图 5.3(b)是总线型结构，分为交叉开关总线系统与传统的共享总线系统。交叉开关总线设计思路是，将一体的交叉总线矩阵划分成小的交叉矩阵，中间通过一条高性能的总线连接。在多核 CPU 中，交叉开关总线主要负责控制多个处理器核与多个缓存段、输入/输出设备之间的数据传输。

图 5.3(c)是互联网络型结构，是一种基于空分交换的矩阵组成的内部网络结构，如

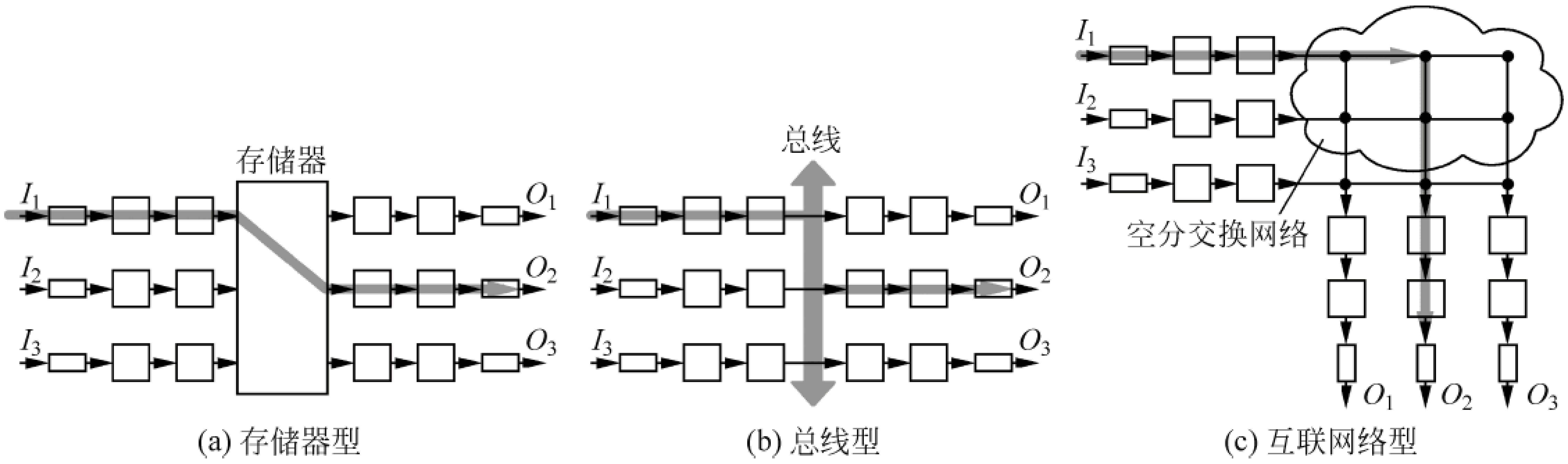


图 5.3　交换结构类型

ATM 交换机就是这种结构，而有些路由器、交换机就是在 ATM 的基础上，配置适当硬件、软件而构成的。

2. 路由交换结构及硬件配置

路由器的硬件配置如图 5.4 所示，各配置解释如下。

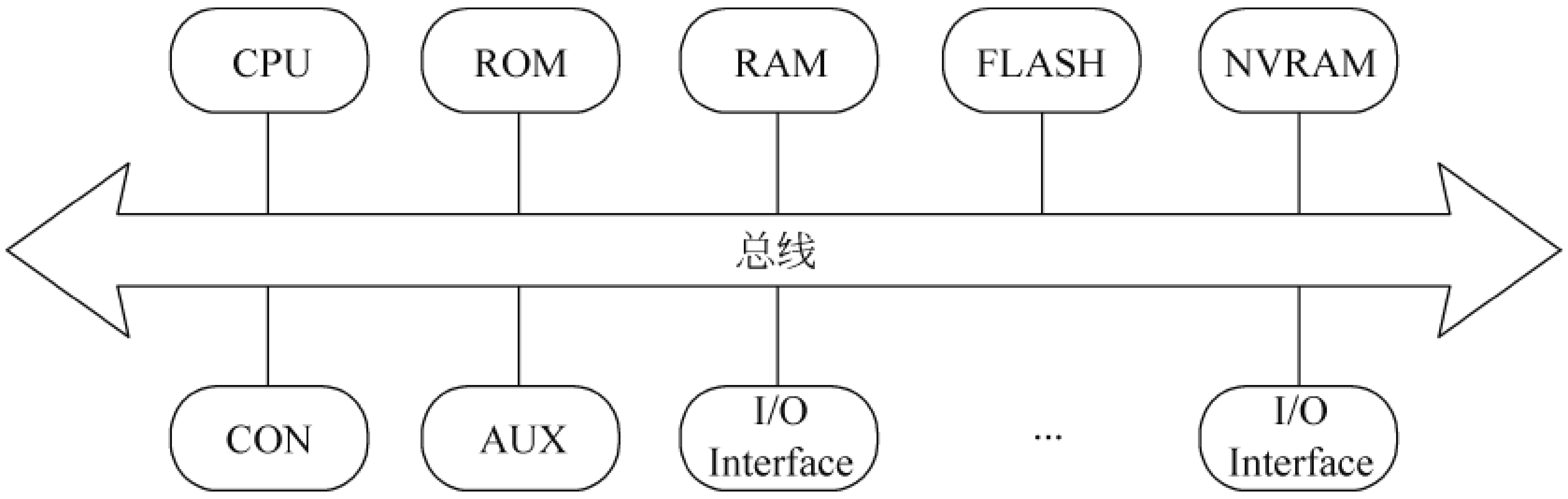


图 5.4　路由器的硬件配置

(1) 中央处理单元(CPU)：作为路由器的中枢，主要负责执行路由器操作系统的指令，以及解释、执行用户输入的命令。同时还完成与计算有关的工作。

(2) 只读存储器(ROM)：只读存储器中包括开机自检程序、系统引导程序及路由器操作系统。

(3) 内存(RAM)：用来存储用户的数据包队列以及路由器在运行过程中产生的中间数据。

(4) 闪存(FLASH)：主要负责保存操作系统的映像文件。

(5) 非易失性内存(NVRAM)：用来存储路由器的启动配置文件。

(6) 控制台接口(CON)：供用户对路由器进行配置使用。

(7) 辅助接口(AUX)：用来连接调制解调器以实现对路由器的远程控制管理。

(8) 接口(Interface)：数据包进出路由器的通道。

(9) 总线(BUS)及缆线：用来连接内部部件以及与其他设备的电缆连接线。

3. 路由器产品

大概分为以下 3 种情况。

(1) 低端路由器：控制、转发都是软件实现。

(2) 中端路由器：控制是软件实现，转发是硬件实现。

(3) 高端路由器：控制部分也基本是软件实现的，转发是硬件实现的，只是性能更高。

5.1.3　路由器工作原理

1. 路由数据包交换

路由器在时刻维持着一张路由表，所有报文的发送和转发都通过查找路由表从相应接口发送。路由器工作流程如图 5.5 所示，可以看出：物理层从路由器的一个接口收到一个报文，上送到数据链路层；数据链路层去掉链路层封装，根据报文的协议域上送到网络层；网络层首先看报文是否是送给本机的，若是，去掉网络层封装，送给上层，若不是，则根据报文的目的地址查找路由表，若找到路由，将报文送给相应接口的数据链路层，数据链路层封装后，发送报文，若找不到路由，则将报文丢弃。

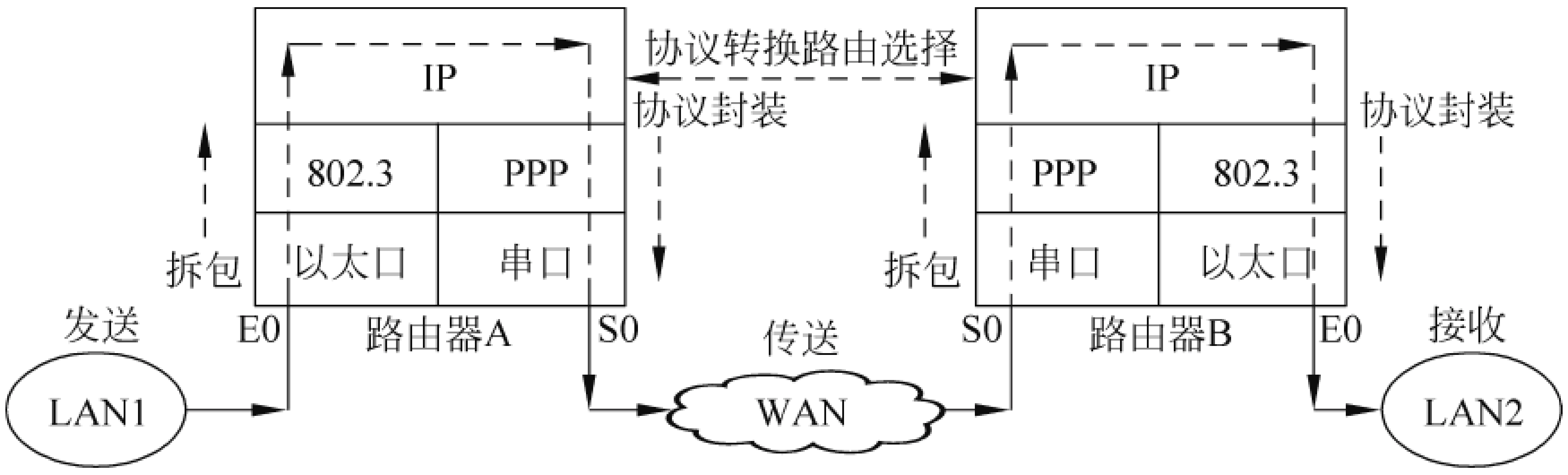


图 5.5　路由器工作流程

一个数据包在被路由的过程中可能要经过若干个路由器节点。其中，每一个节点对该数据包都进行类似的处理。图 5.6 给出了路由过程中的数据包交换原理。图中省略了无关的字段。

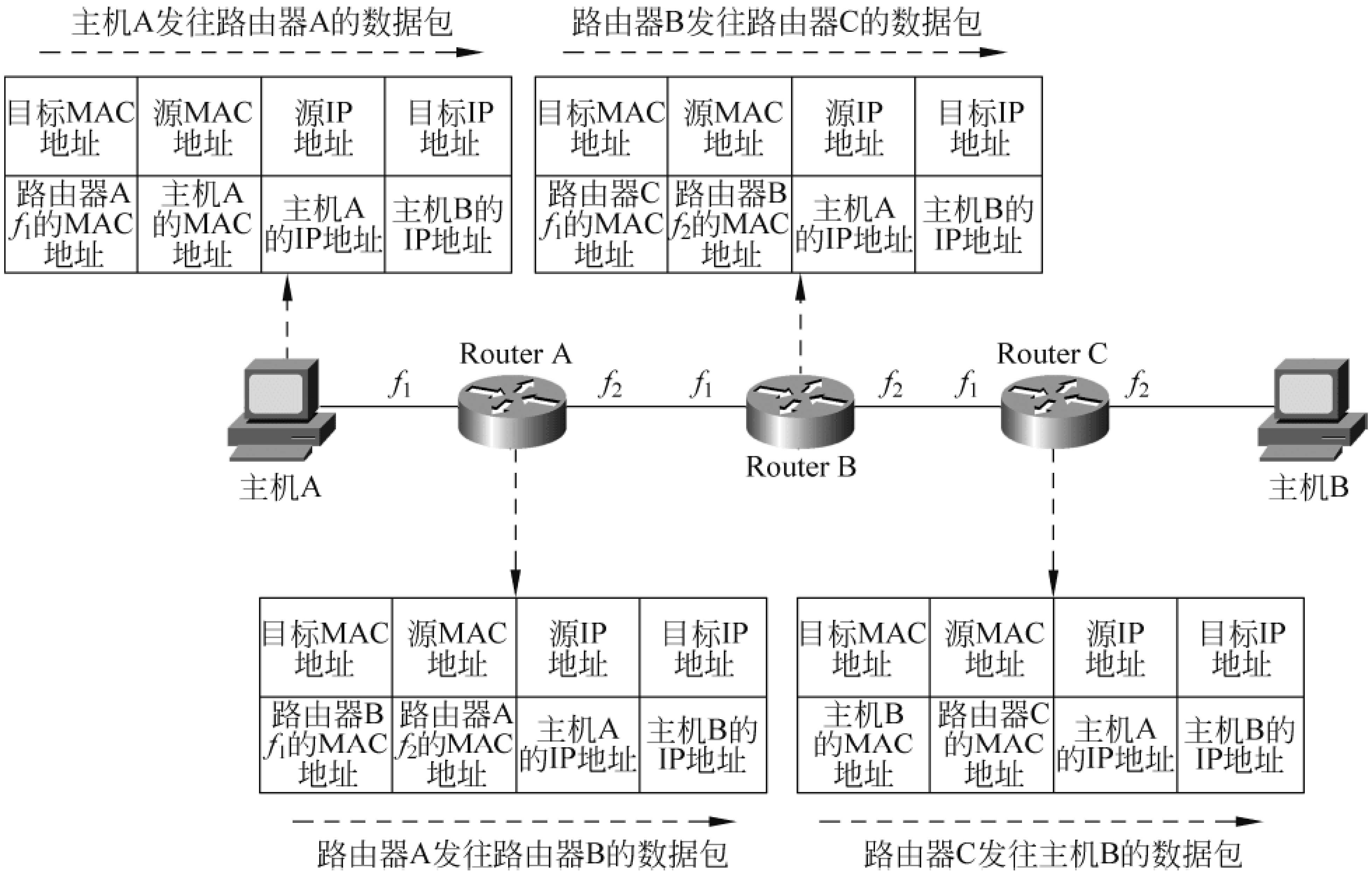


图 5.6　路由过程中的数据包交换示意图

当源节点向位于不同网络上的目标节点发送数据包时，它使用目标节点 IP 地址来发送数据包。在该数据包中，含有本网段路由器的 MAC 地址，通过该 MAC 地址，路由器 A 收

到数据包,然后查看目标节点的 IP 地址。接下来,路由器 A 确定它是否可以转发数据包到目标网络。如果可以转发,该路由器将源 MAC 地址改为自己的 MAC 地址,将目标 MAC 地址改为下一跳设备的 MAC 地址。如果它无法为这个数据包选择路由,则丢弃数据包。

若下一跳不是最终的目标节点,则下一个路由器对数据包执行完全相同的操作,即确定下一跳,更改 MAC 层地址,并转发数据包。如此继续,直到数据包到达目标节点。由此可见,节点 IP 地址一直不会改变,而 MAC 地址在每一跳都要改变。

路由器的寻址转发过程如图 5.7 所示。当 IP 数据报在进入网络层后,路由器软件提取 IP 包里的目标地址,然后与路由表里的目标地址进行逐项比较,先搜索主机匹配项,如找不到匹配的主机时,再找子网;如找不到匹配的子网时,就接着去找有没有相同的网段匹配?如果没有找到匹配项,就将 IP 数据报送上默认路由;如果路由器中没有设置默认路由,就只能将此 IP 包丢弃。

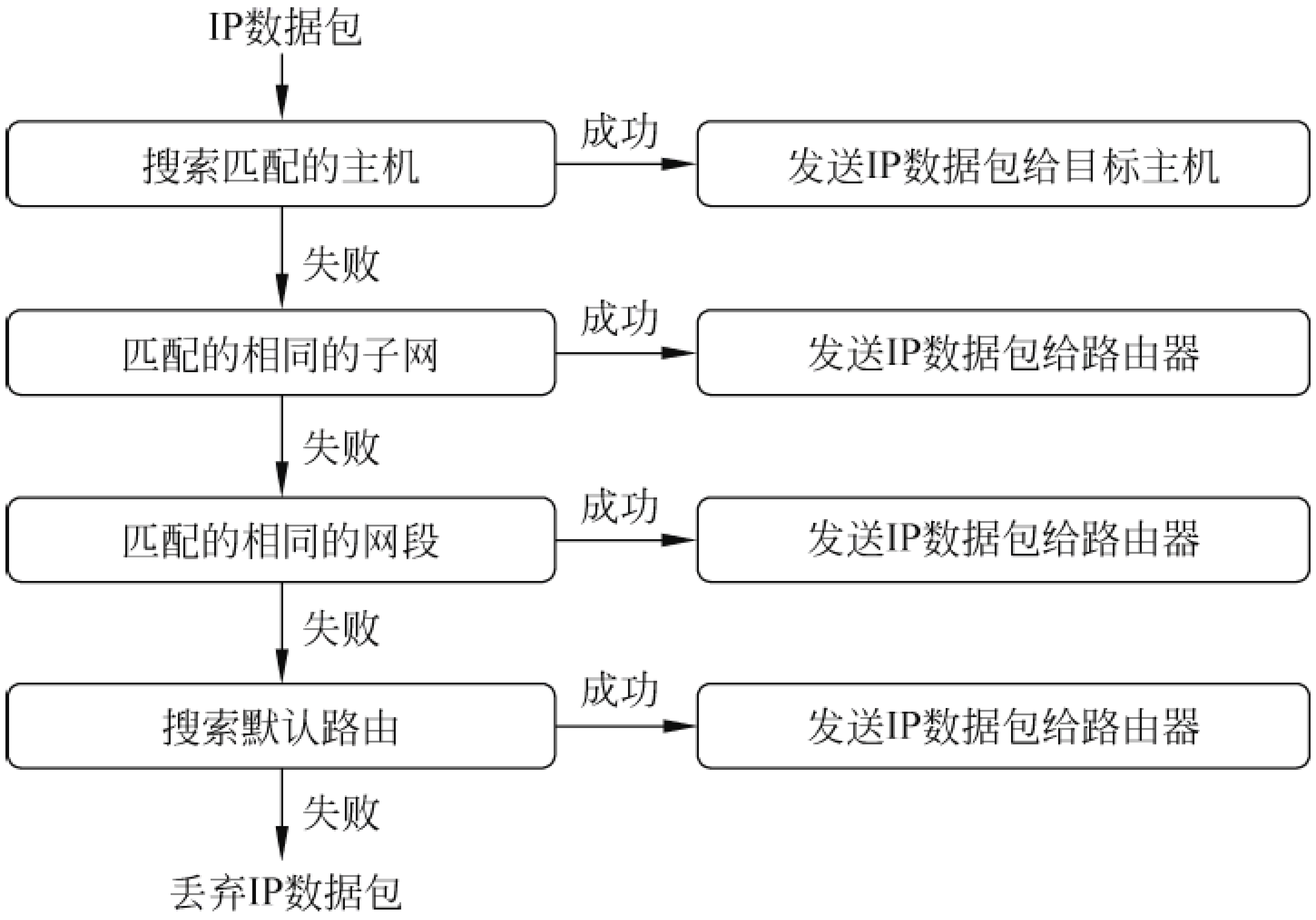


图 5.7 路由器的寻址转发过程

2. 路由器举例说明

【例 1】 已知路由器 R1 的路由表如表 5.2 所示。试根据路由表画出各网络和必要的路由器的连接拓扑,并标注出必要的 IP 地址和接口。

表 5.2 路由器 R1 的路由表

目的网络地址	地 址 掩 码	下一跳地址	路由器接口
150.7.12.64	/24	190.17.0.5	f2
150.7.12.0	/24	190.17.0.5	f2
140.7.8.0	/21	191.17.0.2	f1
120.73.0.0	/16	——	f0
190.17.0.0	/16	——	f2
191.17.0.0	/16	——	f1
0000(默认)	0000(默认)	120.73.0.3	f0

答:根据路由表,绘制的拓扑图如图 5.8 所示,这里要强调的是下一跳地址指的是相邻路由器的接口 IP 地址,路由器接口指的是本路由器自身的相应接口。

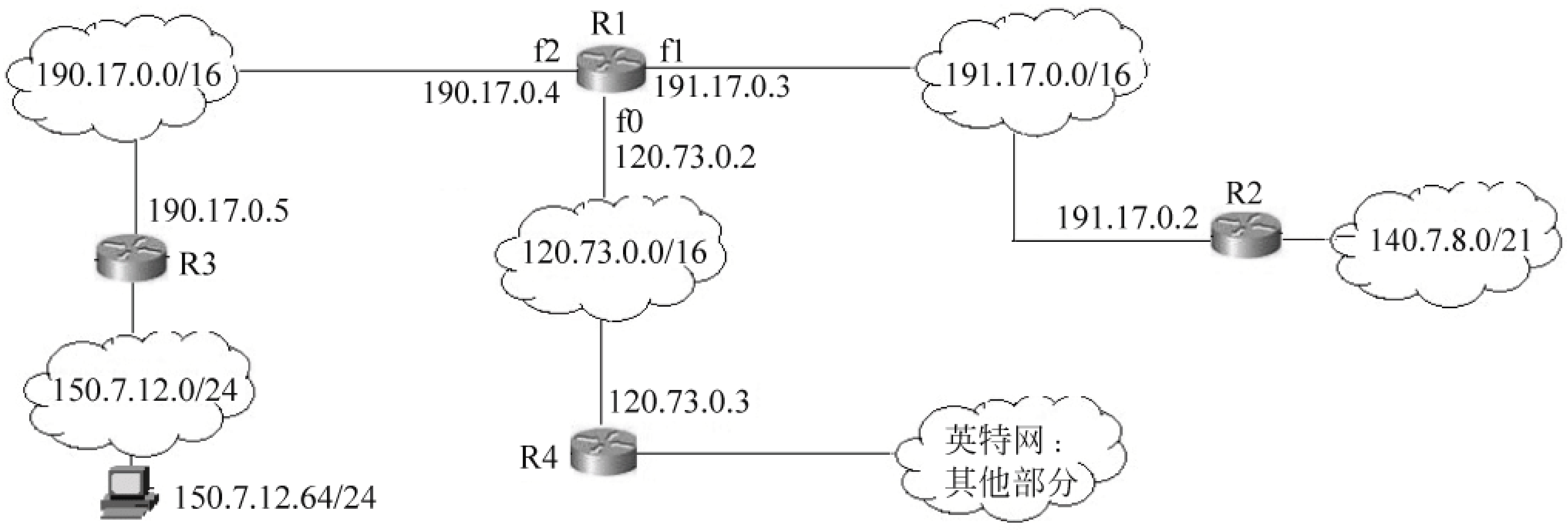


图 5.8　网络拓扑图

【例 2】 某网络中路由器 R1 的路由表信息如表 5.3(a)所示，表中有“目标网络”“距离”“下一跳路由器”的信息，如现在 R1 收到从它的相邻路由器 R3 发来的路由信息，如表 5.3(b)所示，这时，表 5.3(a)的信息就需要更新，试求出路由器 R1 更新后的路由表，并详细说明原因。

表 5.3(a)　路由器 R1 的路由表

目标网络	距　离	下一跳路由器
Net1	4	R2
Net2	2	R3
Net3	1	R4
Net4	5	R3
……		

表 5.3(b)　路由器 R3 发来的更新路由信息

目标网络	距　离	下一跳路由器
Net1	2	R2
Net2	1	R5
Net3	3	R6
Net4	6	R7
Net5	6	R8

答： 这里不必去关心网络的拓扑关系，只要对路由表进行相应的计算就可以得到新的路由表。首先当路由器 R1 收到表 5.3(b)的信息后，路由器 R1 将对 R3 发来的路由信息表 5.3(b)进行处理，因为路由器相邻所以每行的距离要加 1，并将下一跳路由器一律改为提供给它信息的路由器 R3，即产生了修改后的表 5.3(c)。

表 5.3(c)　修改后的表(b)

目标网络	距　离	下一跳路由器
Net1	3	R3
Net2	2	R3
Net3	4	R3
Net4	7	R3
Net5	7	R3

接下来的就是将表 5.3(c)和表 5.3(a)进行合并,产生表路由器 R1 新的路由表,即表 5.3(d)。其理由如下:

第 1 行,不同的下一跳,但新的路由距离更短,因此 R1 距离要改变;

第 2 行,相同的下一跳,距离一样,距离不改变;

第 3 行,不同的下一跳,但新的距离更大,因此 R1 距离不改变;

第 4 行,由于经过 R3 的新路由距离增大了,因此 R1 距离也要改变;

第 5 行,因为表中没有,需要添加这一新项。

表 5.3(d) 路由器 R1 更新后的路由表

目标网络	距离	下一跳路由器
Net1	3	R3
Net2	2	R3
Net3	1	R4
Net4	7	R5
Net5	7	R3

5.2 路由器配置

路由器可以看成是一种用来完成数据包存储、选路和转发的专用计算机。网络设备厂商根据企业网络规模大小及业务应用等的不同特点生产出一系列自成体系的路由器产品家族,以满足市场需要。不同厂商、品牌的路由器产品的配置方法不尽相同。但是其基本工作原理却是相同的,有差异的只是命令字或格式而已。只要了解一种路由器产品的相关配置,就完全可以掌握绝大部分路由器配置的技巧和方法。

5.2.1 路由器启用

1. 路由器启动过程

路由器启动过程如图 5.9 所示,一个新路由器进行第一次加电启动后,会自动运行配置向导。利用配置向导,可以通过问答的形式对路由器进行初始配置。

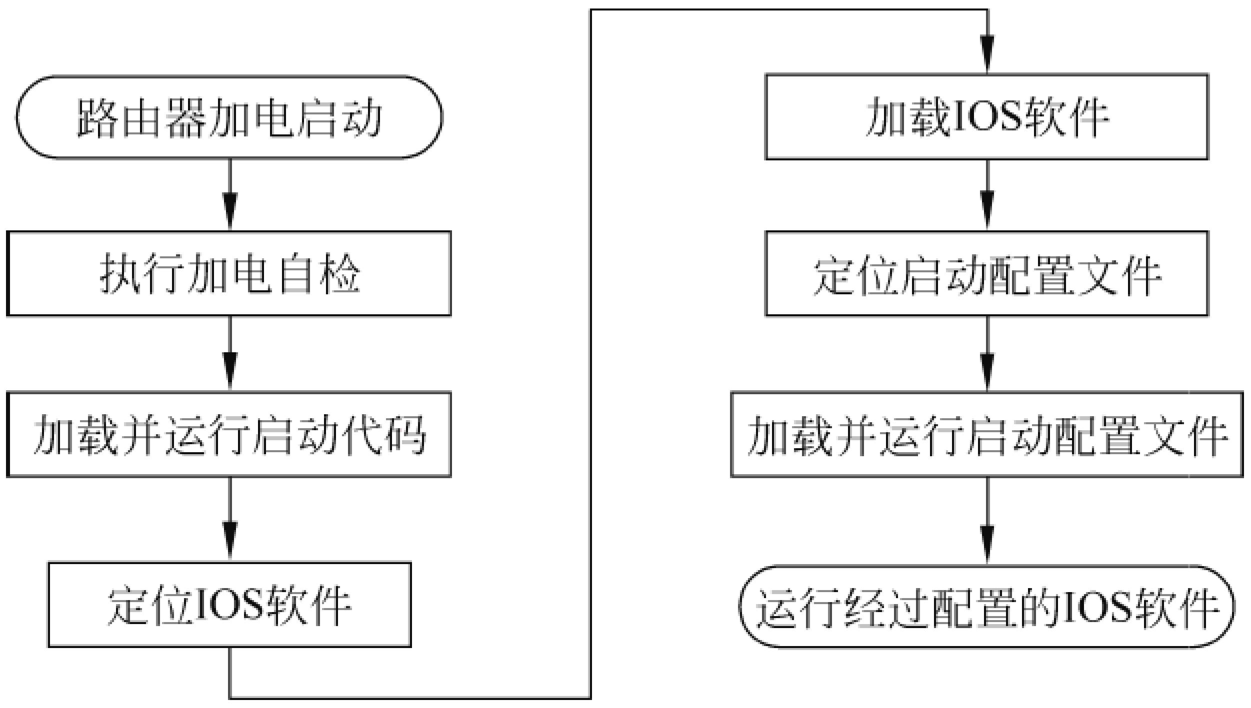


图 5.9 路由器启动过程

图 5.10 所示是对路由器进行多种配置的连接,通常可以使用 5 种方式来设置路由器:控制台接口(Console,CON)连接终端或运行终端仿真软件的 PC,在本地进行配置路由器;利用辅助接口(AUX)连接调制解调器(MODEM),通过电话线与远方的终端或运行终端仿真软件的 PC 相连,对路由器进行远端配置;通过 Ethernet 上的 FTP/TFTP 服务器对路由器进行配置;通过 Ethernet 上的 Telnet 程序登录到路由器进行配置;通过 Ethernet 上的 SNMP 网管工作站,利用网管软件对路由器进行配置。以下简单介绍常用的两种配置方式。

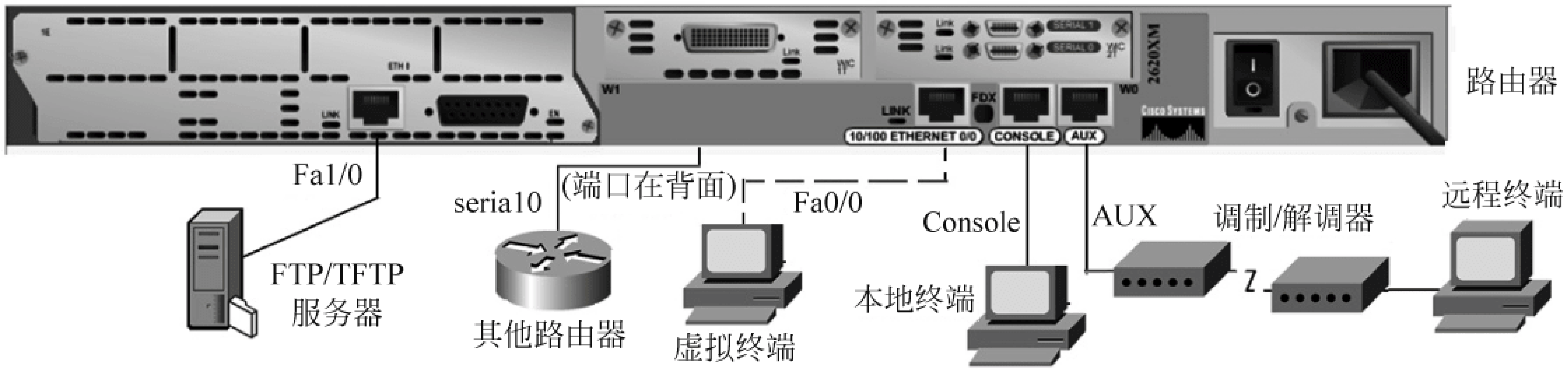


图 5.10 对路由器进行配置的多种连接

1) 通过超级终端进行配置

对于初始安装的路由器来说,只能通过控制台(Console)接口对其进行初始配置。在配置之前,首先必须用路由器附带的控制台电缆连接路由器和终端。正确连接好线缆之后,在工作站端启动"超级终端"应用程序,然后选择连接所使用的接口。最后将终端设备配置成工作于 9600 波特率、8 个数据位、没有奇偶校验、一个停止位的状态。

2) 通过 Telnet 进行配置

当为路由器的某个接口设置了 IP 地址之后,可以通过虚拟终端从任何地点 Telnet 到路由器上对其进行配置。

2. 设置对话过程

一台新路由器开机时自动进入对话状态,而在特权命令状态使用 setup 命令也可进入对话状态,这时可通过对话方式对路由器进行设置。利用设置对话过程可以避免手工输入命令的烦琐,但它还不能完全代替手工设置,一些特殊的设置还必须通过手工输入的方式完成。进入设置对话过程后,路由器显示如下提示信息。

1) 首先进入系统配置对话框(System Configuration Dialog)

```
At any point you may enter a question mark '?' for help
Use ctrl - c to abort configuration dialog at any prompt
Default settings are in square brackets '[ ]'
```

在任何时候,都可以输入一个问号"?"寻求帮助。对话框在任何提示下,按 Ctrl+C 组合键都可以中止配置。默认设置是在方括号"[]"中。接着路由器会问是否进入设置对话:

```
Would you like to enter the initial configuration dialog?
```

如果按 y 或 Enter 键,路由器就会进入设置对话过程。首先可以看到当前的各接口汇总状况:

```
First,would you like to see the current interface summary?
```

你会看到接口名称、IP 地址、工作状态等信息。然后，路由器就开始进入全局参数的设置：

```
Configuring global parameters
```

2）设置路由器名

```
Enter host name [Router]
```

3）设置进入特权状态的密文（secret）

此密文在设置以后不会以明文方式显示。

```
The enable secret is a one - way cryptographic secret used
instead of the enable password when it exists
Enter enable secret:
```

4）设置进入特权状态的密码（password）

此密码只在没有密文时起作用，并且在设置以后会以明文方式显示。

```
Enter enable password:
```

5）设置虚拟终端访问时的密码

```
Enter virtual terminal password:
```

6）询问是否要设置路由器支持的各种网络协议

```
Configure SNMP Network Management?
Configure DECnet?.
Configure AppleTalk?
Configure IPX? [no]:
Configure IP? [yes]:
……
```

7）设置异步接口的参数如果配置的是拨号访问服务器，系统还会设置异步接口的参数。

```
Configure Async lines?
```

（1）设置线路的最高速度。

```
Async line speed [9600]:
```

（2）是否使用硬件流量控制配置。

```
Configure for HW flow control?
```

（3）是否设置 modem。

```
Configure for modems?
```

（4）是否使用默认的 modem 命令。

```
Configure for default chat script?
```

(5) 是否设置异步口的 SLIP/PPP 参数。

```
Configure for Dial-in IP SLIP/PPP access?
```

(6) 是否使用动态 IP 地址。

```
Configure for Dynamic IP addresses?
```

(7) 是否使用默认 IP 地址。

```
Configure Default IP addresses?
```

(8) 是否使用 TCP 头压缩。

```
Configure for TCP Header Compression?
```

(9) 是否在异步口上使用路由表更新。

```
Configure for routing updates on async links?
```

(10) 是否设置异步口上的其他协议。

接下来,系统会对每个接口进行参数设置。

```
Configuring interface Ethernet0:
```

(11) 是否使用此接口。

```
Is this interface in use?
```

(12) 是否设置此接口的 IP 参数。

```
Configure IP on this interface?
```

(13) 设置接口的 IP 地址。

```
IP address for this interface:
```

(14) 设置接口的 IP 子网掩码。

```
subnet mask for this interface:
```

在设置完所有接口的参数后,系统会把整个设置对话过程的结果显示出来。显示结束后,系统会问是否使用这个设置。

```
Use this configuration?
```

如果回答 yes,系统就会把设置的结果存入路由器的 NVRAM 中,然后结束设置对话过程,使路由器开始正常的工作。

5.2.2 路由器配置模式

路由器各个配置模式之间的转换如图 5.11 所示,以下对各种模式进行介绍。

1. (普通)用户 EXEC 模式(user mode)

```
router>
```

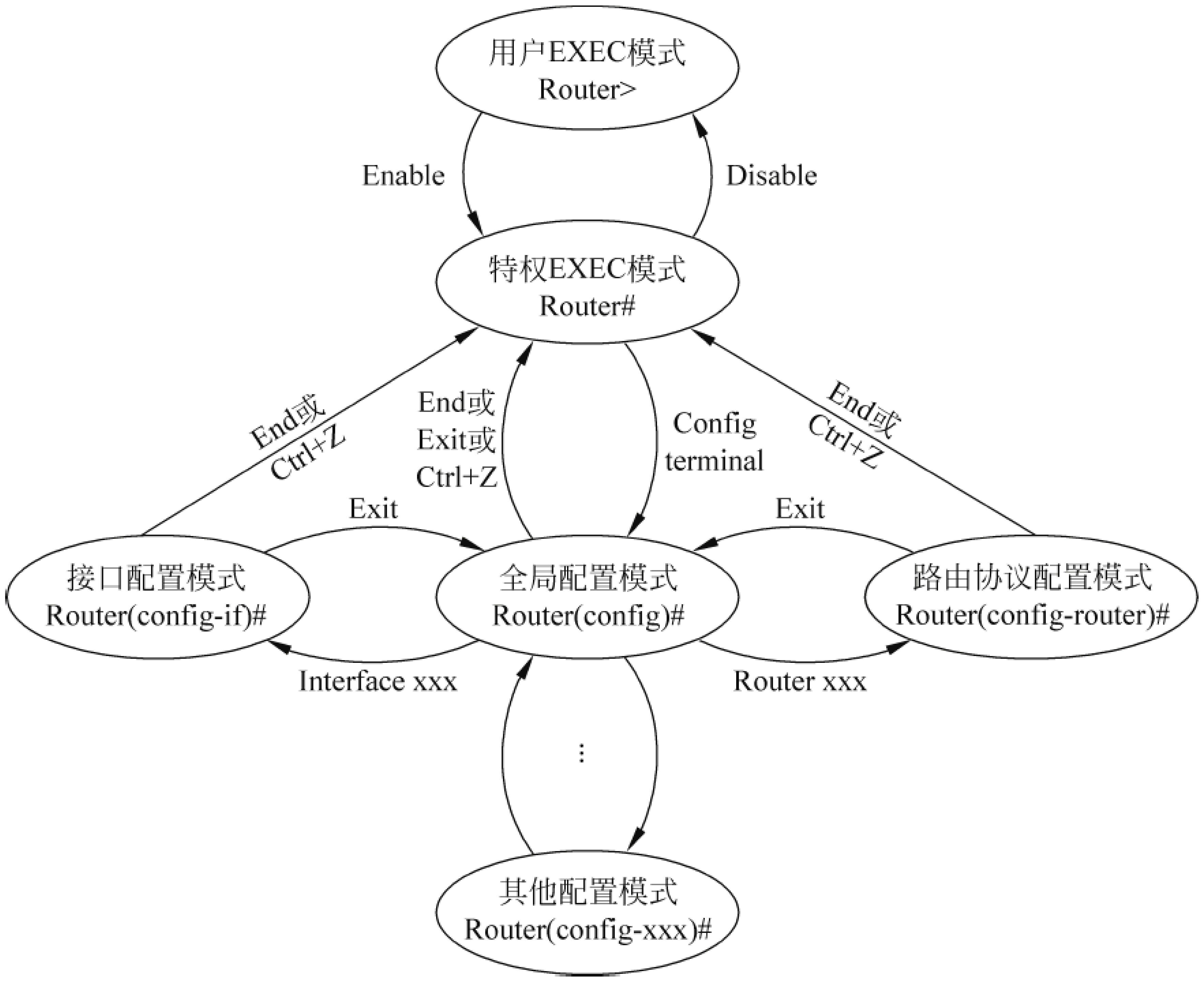


图 5.11　路由器各个配置模式之间的转换

路由器处于用户命令状态，这时用户可以查看路由器的连接状态，访问其他网络和主机，但不能看到和更改路由器的设置内容。

2. 特权用户 EXEC 模式（privileged mode）

```
router＃
```

在 router >提示符下输入 enable，路由器进入特权命令状态 router＃，这时不但可以执行所有的用户命令，还可以看到和更改路由器的设置内容。

3. 全局配置模式（global configuration mode）

```
router(config)＃
```

在 router＃提示符下输入 configure terminal，出现提示符 router(config)＃，此时路由器处于全局设置状态，便可以设置路由器的全局参数。

4. 局部设置状态模式

在 router(config)＃提示符下，输入不同的命令后，路由器就处于局部设置状态，这时可以设置路由器某个局部的参数。常用的几种模式如下。

```
接口配置模式：router(config-if)＃
线命令配置模式：router(config-line)＃
路由协议配置模式：router(config-router)＃
```

5. RXBoot 模式

```
>
```

路由器处于 RXBoot 状态，在开机后 60s 内按 Ctrl＋Break 组合键可进入此状态，这时路由器不能完成正常的功能，只能进行软件升级和手工引导。RXBoot 模式一般在 ROM 中具备基本的 IP 连通性，用于闪存出现故障而用户需要通过 IP 连接来复制一份新的 IOS 到闪存中的情况。该模式通常也可以用于密码丢失，要进行破密进入时。

6. ROMMonitor 模式

```
rommon
```

通常用于 Cisco TAC 的低级调试及口令恢复，当路由器启动时没有找到 IOS 时，自动进入该模式。启动路由器，并在前 60s 内按下 Break 或 Ctrl＋Break 键，即进入 ROM Monitor 模式，也可以通过 PC 的超级终端，使用路由器的 Console 口直接对路由器进行操作。

7. Setup 模式

该模式通常是在配置文件(configuration file)丢失的情况下进入的，以进行手动配置。在此模式下只保存着配置文件的最小子集，再以问答的形式由管理员选择配置。

5.2.3 路由器基本配置

使用路由器向导可以进行基本的配置。对于更详细的参数、选项的配置，只能从互联网操作系统(Internetwork Operating System，IOS)手工完成。

1. 配置基础

1) 路由器命令解释器

缩写：Cisco 路由器允许使用命令的缩写，只要不和别的命令混淆，便可使用尽量短的缩写形式。如：interface 可缩写为 int，如有歧义，系统将给出 Ambiguous command(命令歧义)。

问号：命令列表。在 IOS 操作中，无论任何状态和位置，都可以输入“?”得到系统的帮助。如果忘记命令拼写，可以在输入命令前几个字母后打个问号；如果是参数或子命令，可以在一个命令后加入空格，再输入一个问号。

Tab 键：补全命令。当输入几个命令字母后，再按 Tab 键，后面的命令字母将会被补齐。

2) 命令编辑快捷键

常用的一些命令编辑快捷键见表 5.4。

表 5.4 路由器各个配置模式之间的转换命令编辑快捷键

快 捷 键	作 用	快 捷 键	作 用
BackSpace，Ctrl＋H	删除当前光标左侧的一个字符	Ctrl＋K	删除从光标开始直到行尾的所有字符
Ctrl＋P 或上箭头	重新显示前一命令	Ctrl＋X	删除光标之前的所有字符
Ctrl＋N 或下箭头	重新显示后一命令	Ctrl＋W	删除一个字
Ctrl＋A	到行首	Ctrl＋U	删除一行
Ctrl＋E	到行尾	Ctrl＋R	刷新刚输入的字符
Ctrl＋B	回退一个字符(并不删除)	Esc＋B	回退一个单词
Ctrl＋F	前进一个字符	Esc＋F	前进一个单词
Ctrl＋D	删除光标处的一个字符	Esc＋D	删除光标后的一个单词

3）本书配置命令行语法说明(也适用于交换机命令)

在一条配置命令行的提示符(如＃，＞等)后面，其语法表示形式做如下规定：

粗体字表示照原样输入的命令和关键字。如命令前面有提示符，就不再用粗体字表示；

正常字表示用户应提供的具体参数值。如果在某种模式下的命名行一律用正常字表示；

竖线或管道符(|)用于分隔开可选的、互斥的选项；

方括号([])表示任选项；

花括号({})表示必选项；

方括号中的花括号([{}])表示必须在任选项中选择一个；

空格键( )用于隔离命令字符或参数，以保证每个命令字或参数都是独立的；

横杠(—)用于表明子接口，另外可表示范围或特殊命令字的连接符等；

感叹号(!)表示在它后面的内容为注释，不属于命令内容。只限本书采用，与具体配置的命令注释符是有差异的。

4）搜索、过滤 show 命令的输出结果

show 命令使用频繁，但是有时输出内容也很多，需要一次次地翻页才能找到有用的信息。而人们往往关心输出内容的某一部分，这样就需要用管道符来过滤命令的输出结果。

```
command | {begin | include | exclude} regular - expression
```

command 是带有若干参数的 show 命令；“|”是管道符，用于过滤输出结果；regular-expression 为正则表达式，限定关键词(正则表达式通常被用来检索和/或替换那些符合某个模式的文本内容)；begin 表示包含输出结果中限定关键词的首行开始显示命令的输出；include|exclude 便是包含或不包含限定词的那些行。

5）路由器口令恢复

当路由器的口令被错误修改或忘记时，可以按如下步骤进行操作。

(1) 开机时按 Ctrl＋Break 使进入 ROM 监控状态；

(2) 重新启动路由器：rommon＞reset；

(3) 在 Setup 模式，对所有问题回答 No；

(4) 下载 NVRAM：Router＞configure memory。

2. 改变配置状态等基本命令

(1) 用户模式命令，进入特权命令状态。

```
router＃enable
```

(2) 特权模式命令，退出特权命令状态。

```
router＃disable
```

(3) 特权命令状态下，进入设置对话状态。

```
router＃setup
```

(4) 特权模式，进入全局设置状态。

```
router＃configure terminal
```

(5) 直接退回到特权模。

```
router(config)#end
router(config)#(Ctrl+z)
```

(6) 进入接口设置状态。

```
router(config)#interface type slot / port adaptor / port
```

其中 type 指路由器的接口类型,一般分为 FastEthernet(f)和 Serial(s),其中 f 是以太网接口; s 是串行接口; slot 为接口类型插槽; port adaptor 为接口适配器; port 为接口号。

例,route(config)#interface s2/0　!进入接口 Serial 2/0 设置状态。

(7) 进入子接口设置状态。

```
router(config-if)#interface type number.subinterface[point-to-point | multipoint]
```

其中,point-to-point | multipoint 为点对点|点对多点。

例,route(config-if)#interface s0.1 point-to-point　!设置 S0 的子接口 1,为点对点通信。

(8) 进入线路设置状态。

```
router(config)#line type slot/number
```

说明:type 可以是 Console(控制台接口)、Aux(辅助接口)、Tty(终端控制器接口线路)、Vty(虚拟终端线路)。

(9) 进入路由设置状态。

```
router(config)#router protocol
```

(10) 退出到前一个命令模式。

```
router(config)#exit
```

3. 显示命令

(1) 查看当前设备运行设置。

```
router#show running-config
```

(2) 查看开机(初始配置文件的内容)设置。

```
router#show startup-config
```

(3) 显示路由器硬件配置、软件版本等信息。

```
router#show version
```

(4) 显示配置信息。

```
router#showrun
```

(5) 显示所在接口信息。

```
router#show interface
```

（6）显示指定接口信息。

```
router＃show interface type slot/number
```

（7）显示路由信息。

```
router＃show ip route
```

（8）显示、设置路由器系统时间。

```
router＃show clock/clock set
```

（9）显示历史命令。

```
router＃show history
```

（10）显示邻居信息。

```
router＃show cdp nei
```

4. 路由器基本配置命令

（1）启动 IP 路由。

```
router＃ip routing
```

（2）重新起动路由器。

```
router＃reload
```

（3）配置接口的 IP 地址、子网掩码。

```
router(config－if)＃ip address ip－address subnet－mask
```

（4）设置第二个 IP 地址。

```
router(config－if)＃ip address second
```

（5）激活当前接口。

```
router(config－if)＃no shutdown
```

（6）关闭接口。

```
router(config－if)＃shutdown
```

（7）设置同步时。

```
router(config－if)＃clock rate 64000
```

（8）绑定 VLAN 中继协议。

```
router(config－subif.1)＃encapsulation dot1q
```

（9）修改历史命令。

```
router＃terminal history size
```

（10）保存当前设备运行设置信息。

```
router(config-line)# write memory
```

5. **路由器口令设置**

(1) 设置路由器名。

```
router(config)# hostname username
```

例,router(config)# hostname routerA

(2) 设置特权加密口令。

```
router(config)# enable password userpassword
```

例,route(config)# enable password 654321

(3) 进入控制台口。

```
router(config)# line console 0
```

(4) 进入虚拟终端线路线命令配置模式。

```
router(config)# line vty 0 4
```

设置路由器的远程登录的虚拟端口,0 4 表示可以同时打开 5 个会话,最多为 0 到 15,一共 16 个终端。

(5) 要求口令验证。

```
router(config-line)# login{local | tacacs server}
```

线命令配置模式,设置登录中断时检查密码,启动登录进程。tacacs 为本地用户配置认证和授权;终端访问控制器控制系统协议(Terminal Access Controller Access Control System,TACACS)是一种用于认证的计算机协议,允许远程访问服务器与认证服务器通信,以决定用户是否有权限访问网络。TACACS 允许客户端接受用户名和口令,TACACS 服务器(tacacs server)一般是在主机上运行的一个程序。

(6) 设置登录口令。

```
router(config-line)# password loginpassword
```

(7) 查询并记录丢失的口令。

```
router# show configuration (show startup-config)
```

(8) 设置路由器加密使能密码。

```
route(config)# enable secret password
```

例,route(config)# enable secret 123456

6. **路由器文件管理操作命令**

路由器有两种配置文件 running-config(当前配置文件)和 startup-config(初始配置文件)。用 show 命令可以分别查看两个文件的内容。参考图 5.12,下面是常见的几种命令。

(1) 当前配置文件复制到初始配置文件中,保存到 Flash。

```
route# copy running-config startup-config
```

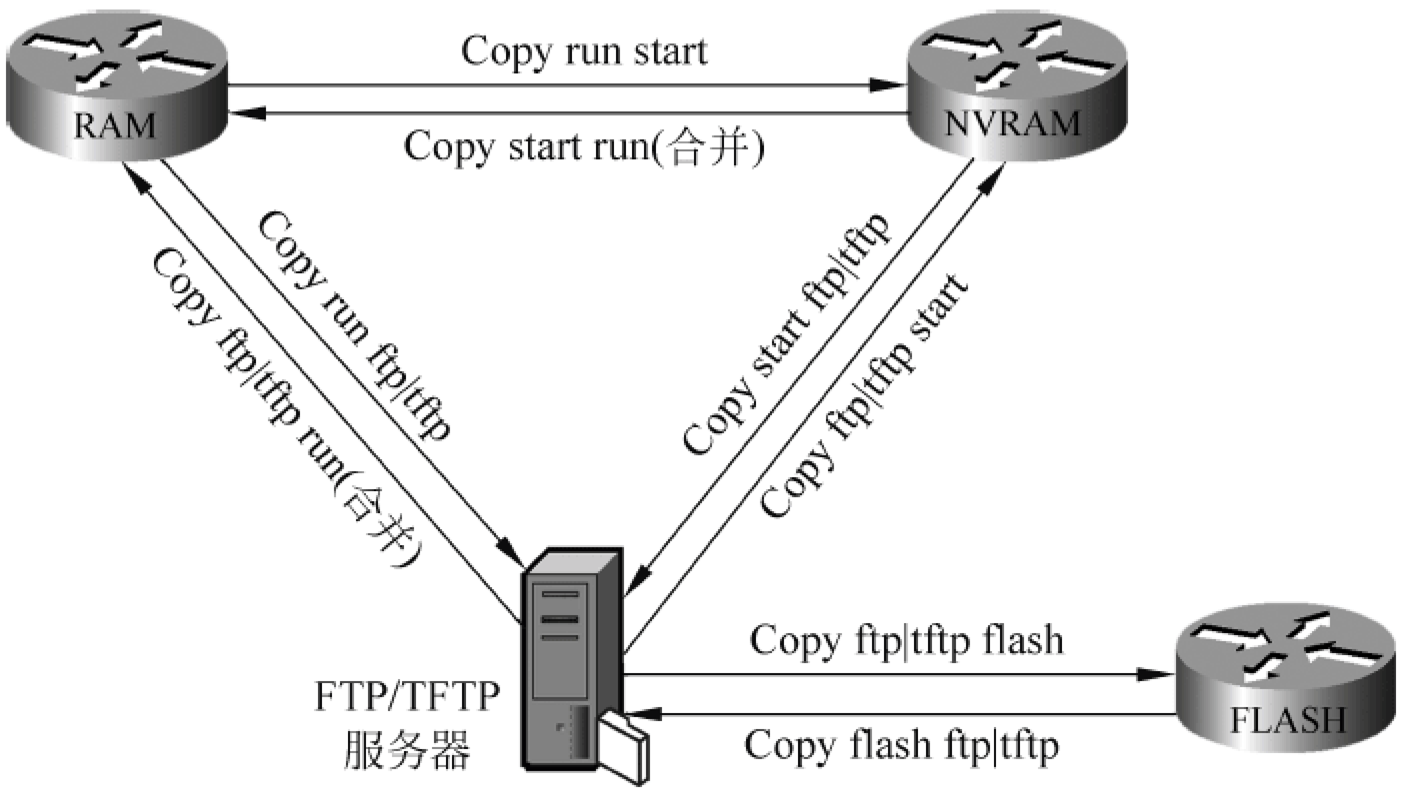


图 5.12 常见的几种复制方式及其命令

(2) 初始配置文件复制到当前配置文件。

```
router#copy startup-config running-config
```

(3) 复制运行配置到 TFTP。

```
router#copy startup-config tftp
```

(4) 开机配置存到 TFTP。

```
router#copy startup-config tftp
```

(5) 下传文件到 Flash。

```
router#copy tftp flash
```

(6) 下载配置文件。

```
router#copy tftp startup-config
```

(7) 删除启动配置文件。

```
router#erase startup-config
```

(8) 进入 ROM 监视器模式“rommon>”。

```
Ctrl+Break
```

(9) 跳过配置文件。

```
rommon>confreg 0x2142
```

(10) 恢复配置文件。

```
rommon>confreg 0x2102
```

(11) 重新引导。

```
rommon>reset
```

(12) 从 console 传输文件。

```
rommon>copy xmodem flash
```

（13）在虚拟终端登录窗口显示路由、系统和调试信息。

```
terminal monitor
```

（14）从 TFTP 下载。

```
rommon> tftp dnld
```

（15）查看闪存内容。

```
rommon> dir flash
```

（16）引导 IOS。

```
rommon> boot
```

7. 网络命令

（1）登录远程主机。

```
telnet hostname | IP address
```

（2）网络侦测。

```
ping hostname | IP address
```

普通、特权模式命令，用于查看路由信息。

（3）路由跟踪。

```
trace hostname | IP address
```

8. 路由器命令的编辑功能

（1）全局禁用编辑模式。

```
route(config-line)#no editing
```

（2）当前终端线路重新启用编辑模式。

```
router#terminal editing
```

（3）线路配置编辑模式。

```
route(config-line)#editing
```

5.2.4 静态路由的配置

在组网结构比较简单的网络中，只需配置静态路由就可以使路由器正常工作，仔细设置和使用静态路由可以改进网络的性能，并可为重要的应用保证带宽。还有一种接口静态路由，它用于表示那些直接连接到路由器接口上的目的网络。接口静态路由优先级是 0，这意味着它是直接连接网络的路由。

1. 静态路由属性与命令

静态路由属性如下。

可达路由：正常的路由都属于这种情况，即 IP 报文按照目的地标示的路由被送往下一

跳，这是静态路由的一般用法。

目的地不可达的路由：当到某一目的地的静态路由具有 reject（拒绝）属性时，任何去往该目的地的 IP 报文都将被丢弃，并且通过 ICMP 消息通知源主机目的地不可达。

目的地为黑洞的路由：当到某一目的地的静态路由具有 blackhole（黑洞）属性时，任何去往该目的地的 IP 报文都将被丢弃。与 reject 的区别是不向源主机发送任何消息。

1）Quidway 路由器命令

```
ip route <ip_address> [<mask>|<masklen>] <interface_name>|<gateway_address>
[preference <preference_value>] [ reject|blackhole]
```

其中各参数的解释如下。

（1）<ip_address>[<mask>|<masklen>]：目的 IP 地址和掩码。也可用掩码长度表示掩码。

（2）<interface_name>|<gateway_address>：发送接口或下一跳地址。

在配置静态路由时，可指定发送接口 interface_name，也可指定下一跳地址 gateway_address，是指定发送接口还是指定下一跳地址要视具体情况而定。

只有下一跳所属的接口是点到点（PPP、HDLC）的接口时，才能填写<interface_name>，否则，必须填写<gateway_address>。

实际上，所有的路由项都必须明确下一跳地址。IP 在发送报文时，首先根据报文的目的地址寻找路由表中与之匹配的路由。只有路由指定了下一跳地址，链路层才能通过下一跳 IP 地址找到对应的链路层地址，然后按照该地址将报文转发。

在以下几种情况下，有的可以指定发送接口，有的不可以指定发送接口：

对于支持网络地址到链路层地址解析的接口，当 ip-address 和 mask（或 mask-length）指定了一个主机地址，而且该目的地址就在该接口的直接连接网络中，这时可以指定发送接口；

对于点到点接口，指定发送接口即隐含指定了下一跳地址，可以认为与该接口相连的对端接口地址就是路由的下一跳地址，如串口封装 PPP 协议；

对于 NBMA（非广播多路访问网络）接口（如封装 X. 25 或帧中继的接口等），支持点到多点，这时除了配置 IP 路由外，还需要在链路层建立二次路由，即 IP 地址到链路层地址的映射（如 dialer map ip、x. 25 map ip 或 frame-relay map ip 等）。这种情况下配置静态路由就不能指定发送接口，而应配置下一跳 IP 地址。

（3）<preference_value>：优先级。

对优先级 preference 的不同配置，可以灵活应用路由管理策略。

（4）其他参数。

属性 reject 和 blackhole 分别指明不可达路由和黑洞路由。

在网络结构比较简单，且一般到达某一网络所经过的路径唯一的情况下采用静态路由。

2）Cisco 路由器静态路由的配置

```
ip route prefix mask {address | interface} [distance] [permanent]
```

其中：

prefix：所要到达的目的网络；

mask：子网掩码；

address：下一个跳的 IP 地址，即相邻路由器的接口地址；

interface：本地网络接口；

distance：管理距离（可选）；

permanent：指定此路由，即使该接口关掉也不被移掉。

例如：配置访问目标网 192.1.0.64/26，下一跳地址为相邻路由器的 192.200.10.6。

```
routerA(config)# ip route 192.1.0.64 255.255.255.192 192.200.10.6
```

例如：配置访问目标网 192.1.0.128/26，经过直连下一跳路由器的本路由器接口为 s0。

```
routerA(config)# ip route 192.1.0.128 255.255.255.192 s0
```

【例 3】 如图 5.13 所示，在路由器 Router-A 上配置一条到目的网段 129.200.0.0/16 的静态路由，并已知下一跳地址为路由器 Router-B 的 s0 接口的 IP 地址 192.200.0.2。如果链路的封装是 PPP 或 HDLC，也可以指定本路由器的转发接口。

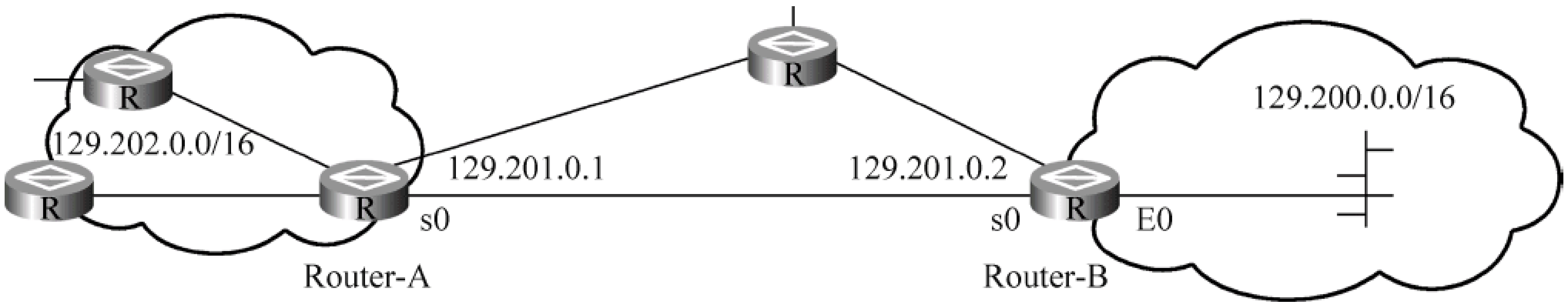


图 5.13　路由器 Router-A 上静态配置网络结构

答：静态路由配置命令如下：

```
routerA(config)# ip route 129.200.0.0 16 s 0
或 routerA(config)# ip route 129.200.0.0 16 192.201.0.2
或 routerA(config)# ip route 129.200.0.0 255.255.0.0 192.201.0.2
```

用户通过配置静态路由，可以人为地指定对某一网络访问时所要经过的路径，在网络结构比较简单，且一般到达某一网络所经过的路径唯一的情况下采用静态路由。

2. 默认路由

参考图 5.13，Router A 的默认路由的配置命令如下。

```
routerA(config)# ip route 0.0.0.0 0.0.0.0 192.201.0.2
```

只要没有在路由表里找到去特定目的地址的路径，则数据均被路由到地址为 192.201.0.2 的相邻路由器。

默认路由也是一种静态路由。简单地说，默认路由就是在没有找到匹配的路由表入口项时才使用的路由。即只有当没有合适的路由时，默认路由才被使用。在路由表中，默认路由以到网络 0.0.0.0（掩码为 0.0.0.0）的路由形式出现。可通过命令 display ip route 的输出看它是否被设置。如果报文的目的地址不能与路由表的任何入口项相匹配，那么该报文将选取默认路由；如果没有默认路由且报文的目的地址不在路由表中，那么该报文被丢弃的同时，将返回源端一个 ICMP 报文指出该目的地址或网络不可达。

默认路由在网络中是非常有用的。在一个包含上百个路由器的典型网络中,选择动态路由协议可能耗费较大量的带宽资源,使用默认路由意味着采用适当带宽的链路来替代高带宽的链路,以满足大量用户通信的需求。

Internet 上绝大部分路由器上都存在一条默认路由。

默认路由并不一定都是手工配置的静态路由,有时也可以由动态路由协议产生。比如 OSPF 路由协议配置了 Stub 区域的路由器会动态产生一条默认路由。

3. 路由自环

参考图 5.12,如果 Router A、Router B 进行了如下配置就构成了路由自环。

```
routerA(config)# ip route 129.204.0.0 255.255.0.0 192.201.0.2
routerB(config)# ip route 129.204.0.0 255.255.0.0 192.201.0.1
```

“路由自环”是指某个报文从一台路由器发出,经过几次转发之后又回到初始的路由器。原因是其中部分路由器的路由表出现错误。产生错误的原因可能是配置静态路由有误,或者是动态路由协议错误地计算路由(虽然这种情况发生的概率很小)。当产生路由自环时,报文会在几个路由器之间循环转发,直至 TTL=0 时才被丢弃,极大地浪费了网络资源,因此应该尽量避免“路由自环”的产生。

5.3 路由器配置实例

本节给出一个实例,网络拓扑如图 5.14 所示,要求写出各个网段地址,然后对各个网络设备进行配置,并验证其配置的正确性及查看有关路由信息。

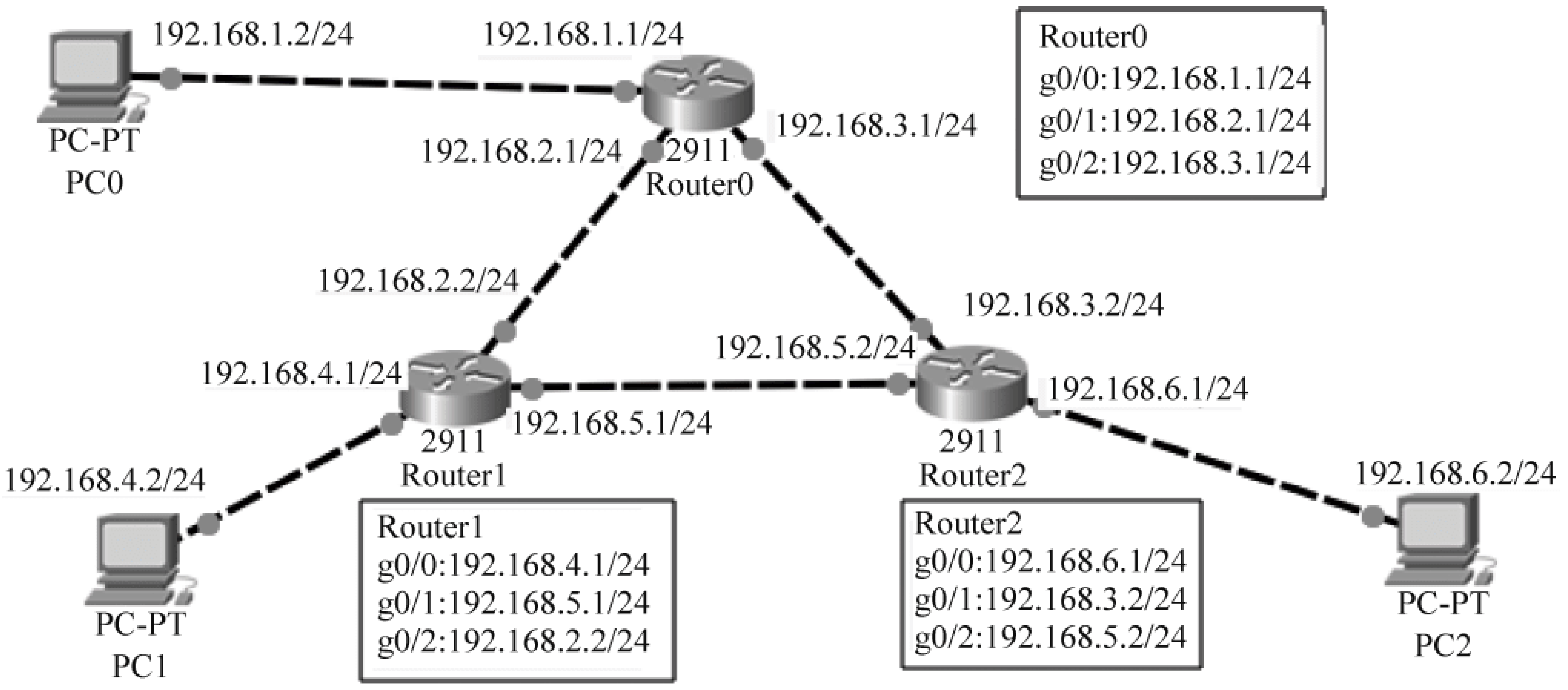


图 5.14 多路由器组网

配置说明:GigabitEthernet0/0、GigabitEthernet0/1、GigabitEthernet0/2 用 g0/0、g0/1、g0/2 代替;Router0 改名为 R0,Router1 改名为 R1,Router2 改名为 R2。

5.3.1 多路由器组网及配置

1. 找出各个网段地址

根据本网络拓扑结构,共找出 6 个网段地址,分别为:

192.168.1.0/24(对应于 R0 的 GigabitEthernet0/0)；

192.168.2.0/24(对应于 R0 的 GigabitEthernet0/1 和 R1 的 GigabitEthernet0/2 直连)；

192.168.3.0/24(对应于 R0 的 GigabitEthernet0/2 和 R2 的 GigabitEthernet0/1 直连)；

192.168.4.0/24(对应于 R1 的 GigabitEthernet0/0)；

192.168.5.0/24(对应于 R1 的 GigabitEthernet0/1 和 R2 的 GigabitEthernet0/2 直连)；

192.168.6.0/24(对应于 R2 的 GigabitEthernet0/0)。

2．对 R0 进行配置

```
Router>enable
Router#configure terminal
Router(config)#hostname R0                                  !改名为 R0
R0(config)#interface g0/0                                   !进入 g0/0 接口配置
R0(config-if)#ip address 192.168.1.1 255.255.255.0          !配置接口 g0/0 对应 IP 地址
R0(config-if)#no shutdown                                   !激活接口 g0/0
R0(config-if)#exit                                          !退回到全局配置模式
R0(config)#interface g0/1                                   !进入 g0/1 接口配置
R0(config-if)#ip address 192.168.2.1 255.255.255.0          !配置接口 g0/1 对应 IP 地址
R0(config-if)#no shutdown                                   !激活接口 g0/1
R0(config-if)#exit
R0(config)#interface g0/2                                   !进入 g0/2 接口配置
R0(config-if)#ip address 192.168.3.1 255.255.255.0          !配置接口 g0/2 对应 IP 地址
R0(config-if)#no shutdown                                   !激活接口 g0/2
R0(config)#ip route 192.168.4.0 255.255.255.0 192.168.2.2   !配置到达 192.168.4.0/24 路由
R0(config)#ip route 192.168.6.0 255.255.255.0 192.168.3.2   !配置到达 192.168.6.0/24 路由
```

3．对 R1 进行配置

```
Router>enable
Router#configure terminal
Router(config)#hostname R1                                  !改名为 R1
R1(config)#interface g0/0
R1(config-if)#ip address 192.168.4.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface g0/1
R1(config-if)#ip address 192.168.5.1 255.255.255.0
R1(config-if)#no shutdown
R1(config-if)#exit
R1(config)#interface g0/2
R1(config-if)#ip address 192.168.2.2 255.255.255.0
R1(config-if)#no shutdown
R1(config)#ip route 192.168.1.0 255.255.255.0 192.168.2.1   !配置到达 192.168.1.0/24 路由
R1(config)#ip route 192.168.6.0 255.255.255.0 192.168.5.2   !配置到达 192.168.6.0/24 路由
```

4．对 R2 进行配置

```
Router>enable
Router#configure terminal
Router(config)#hostname R2                                  !改名为 R2
R2(config)#interface g0/0
```

```
R2(config-if)# ip address 192.168.6.1 255.255.255.0
R2(config-if)# no shutdown
R2(config-if)# exit
R2(config)# interface g0/1
R2(config-if)# ip address 192.168.3.2 255.255.255.0
R2(config-if)# no shutdown
R2(config-if)# exit
R2(config)# interface g0/2
R2(config-if)# ip address 192.168.5.2 255.255.255.0
R2(config-if)# no shutdown
R2(config-if)# exit

R2(config)# ip route 192.168.1.0 255.255.255.0 192.168.3.1   !配置到达 192.168.1.0/24 路由
R2(config)# ip route 192.168.4.0 255.255.255.0 192.168.5.1   !配置到达 192.168.4.0/24 路由
```

5.3.2 测试路由器配置

1. ping 测试

对每个 PC 配置合适的 IP 地址、掩码和网关。如在 PC0 上按截图 5.15 键入有关选项，同样对 PC1、PC2 也做类似的配置。

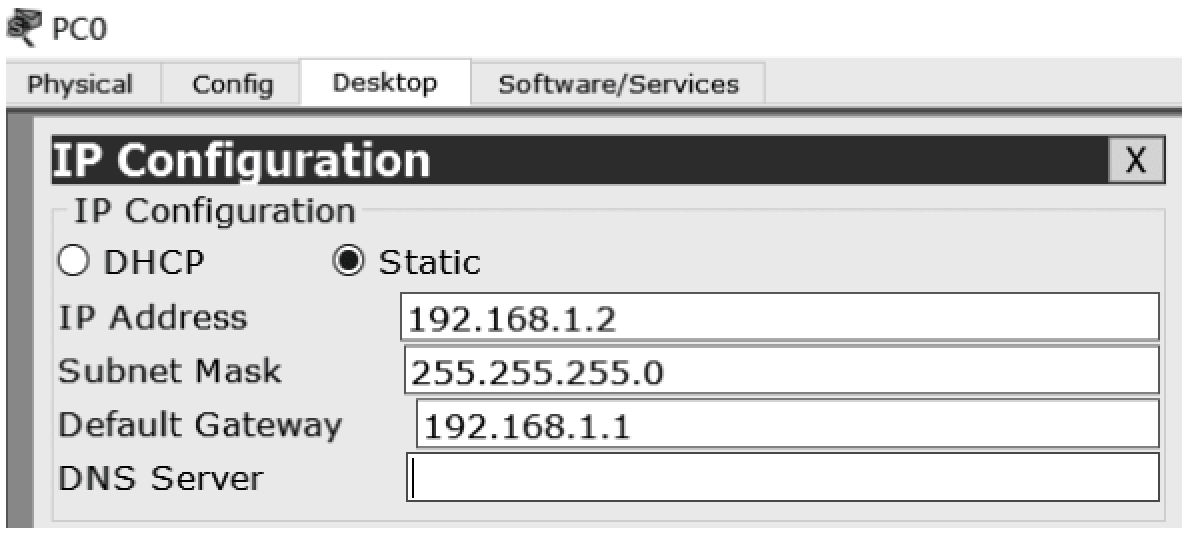


图 5.15 IP 地址配置截图

接着，在 PC 上打开“运行”对话框，在输入框中键入命令 CMD，并单击“确定”按钮即可进入 MSDOS 界面。然后进行 ping 链路测试，其截图如图 5.16 所示。

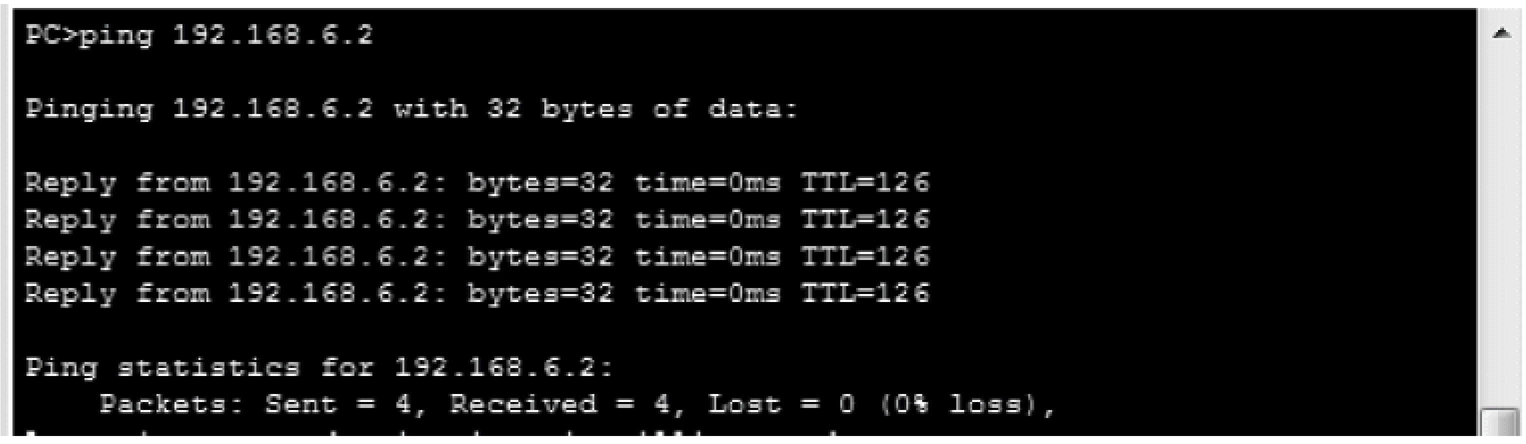


图 5.16 ping 命令截图

从图 5.16 中可以看到，正在 ping 192.168.6.2，具有 32 字节的数据；192.168.6.2 的 ping 统计信息为数据包：已发送=4，已接收=4，丢失=0(0%丢失)等。

在模拟器上可以在 PC 或路由器上，通过对不同目标地发单包或多包，可方便地测试路由配置的成功或失败。图 5.17 给出了模拟器测试表截图，可以看见数据包的源、目的地、协议类型，以及所用时间等信息。

PDU List Window

Fire	Last Status	Source	Destination	Type	Color	Time(sec)	Periodic	Num	Edit	Delete
●	Successful	PC1	PC2	ICMP		0.000	N	0	(edit)	
●	Successful	PC0	Router2	ICMP		0.000	N	1	(edit)	
●	Successful	Router2	PC1	ICMP		0.000	N	2	(edit)	
●	Successful	PC2	Router0	ICMP		0.000	N	3	(edit)	
●	Successful	PC0	192.168.4.2	ICMP		1.000	N	4	(edit)	
●	Successful	PC2	192.168.1.2	ICMP		1.000	N	5	(edit)	
●	Successful	Router2	192.168.4.1	ICMP		1.000	N	6	(edit)	
●	Successful	PC1	192.168.1.1	ICMP		1.000	N	7	(edit)	

图 5.17　模拟器测试表截图

2. 显示 IP 接口

显示 IP 接口截图如图 5.18 所示。在提示符 R1>下，输入 show ip interface 命令时，系统就会告诉你此路由器接口的相关信息：GigabitEthernet0/0 是 up，在线协议是处于 up(连接)；网络地址是 192.168.4.1/24；广播地址 255.255.255.255；通过设置命令确定地址；MTU 是 1500 字节等。

```
R1>show ip interface
GigabitEthernet0/0 is up, line protocol is up (connected)
  Internet address is 192.168.4.1/24
  Broadcast address is 255.255.255.255
  Address determined by setup command
  MTU is 1500 bytes
```

图 5.18　显示 IP 接口截图

3. 查看路由器配置静态路由

通过 R0#show ip route 命令可以查看路由表的配置情况，前面的例题中给每个路由器配置静态路由，才使得每个主机之间能通信。

如果在 R0 上取消已配置的静态路由，以及取消默认路由，如：

```
R0(config)#no ip route 0.0.0.0 0.0.0.0 12.1.3.2
```

这时再用 pc0 ping pc2 和 pc3，就出现问题，网络不通了。

习题

1. 路由是什么？都有哪四类路由？
2. 什么是路由器？并说明路由器的作用。
3. 路由器属于第几层设备？是怎么转发数据包的？
4. 根据路由器的结构，说明路由器的数据包交换原理。
5. 说明路由器启动设置时的对话过程。
6. 简述路由器各个配置模式之间的转换。
7. 简述路由器主要静态路由的配置命令。
8. “路由器是第三层即网络层的设备。根据包的目的 IP 地址来转发，路由器中有一张路由表也叫转发数据库，保存了到网段的最佳路由。匹配路由时是按最长匹配原则来转发包。”你是怎样理解这句话的？路由器是不是最高只能工作在网络层？为什么？
9. 结合并理解图 5.14，用多个路由器组网，并完成相关配置及查看配置结果。

第6章 交换机

交换机(Switch)作为数据交换设备,主要应用于局域网。特别是随着局域网覆盖范围的扩大,只要是没有广域网连接要求,同时又需要路由器的地方,都可以用物美价廉的三层交换机代替。本章主要介绍交换机的基本原理、配置,以及结合VLAN技术的有关应用。

6.1 以太网交换机

在以太网中,交换机仅根据MAC地址进行帧的选路和转发,当一个完整正确的以太网数据帧从一个交换机端口(这里讲的端口都指物理端口)上被接收上来以后,交换机将在自己维护的MAC地址表中去查找地址,根据目标地址类型的不同和查找结果的不同进行交换。

6.1.1 交换技术

1. 网桥的交换原理

网桥是交换机的前身,网桥的交换原理如图6.1所示,网桥A将原来由物理层设备集线器构成的一个总线型局域网,连接并分割成了两个小网段:网段1和网段2。因为网桥A通过端口1和端口2分别连接网段1和网段2,所以网桥会收到各自网段的数据包。当网桥A最初启动后,将自动收到数据包中的源主机MAC地址以及其对应的端口号,并将其存储到缓存表中,这张表称为桥接表或交换表。

网桥A在经过一段时间运行后,它将会学习到网段1和网段2上所有主机的MAC地址以及所在端口。这时网桥A就开始进行转发或过滤,此时,如果网桥A收到一个主机A发送给主机B的单播数据包,网桥A查找自己的桥接表,发现主机B和主机A在同一个端口,则网桥A将执行过滤功能,丢弃此数据包。相反,如果网桥A收到一个主机C发送给主机X的单播数据包,网桥A查找桥接表,发现主机C和主机X处于不同的端口,则网桥A将执行转发功能,将此数据包转发到端口2所在的网段2。

网桥工作在第二层,与集线器相比具有智能性、自学习功能,集线器连接的主机同处于一个大的碰撞域,而网桥的功能就是减小碰撞域,将一个碰撞域分为两个网段,就形成了两个碰撞域,每个网段为一个碰撞域,有效地减少了以太网内主机的碰撞概率。但无论集线器,还是网桥所连接的主机仍然同处于一个广播域中,不能避免广播风暴的袭击。

在以太网中,所谓广播域就是指在一个网络中,广播帧(目的MAC地址为ff-ff-ff-ff-ff-ff的帧)将要被转发的最大范围。当出现广播帧或在地址表中不能匹配到目的地址的帧时,帧将被广播到整个网络上,全部的主机都将接收到。由于网络广播和目的地址未匹配的帧的普遍存在,当二层网络的规模增加,网络广播严重,网络运行效率下降,也不利于网络的有效管理。

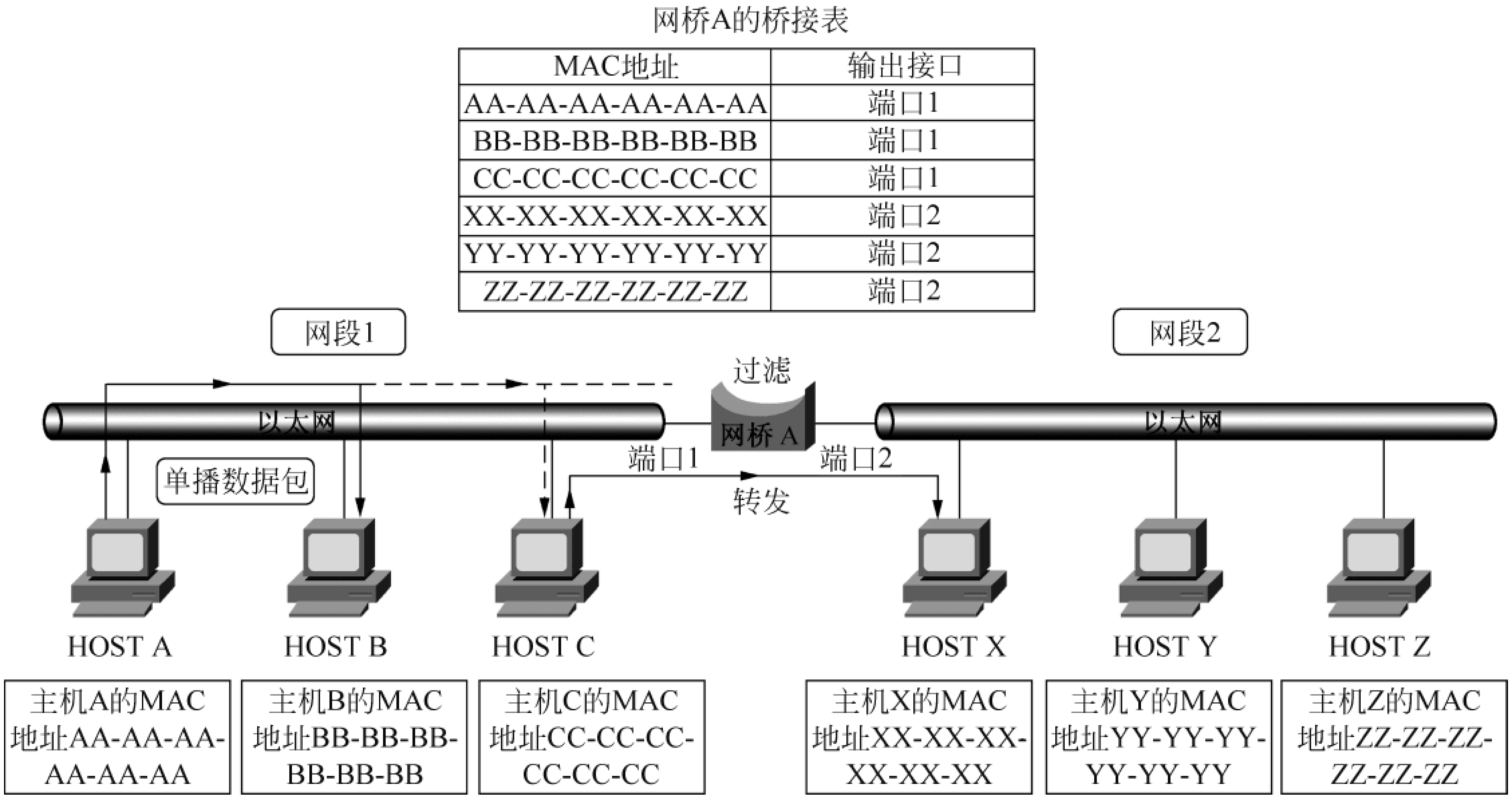


图 6.1　网桥的交换原理示意图

由于整个网络在一个广播域,所有的用户都能够不受控制地直接访问网络的所有部分,并能够影响到网络所有部分的正常运行,因此对于网络的安全性也造成一定的威胁。

可以看出,网桥的操作数据包实际上就是在第二层上的数据帧(frame),数据帧的所谓"交换"实际上就是指转发,需要执行两个基本的操作过程:第一,地址学习过程。用以构造和维护桥接表,以便交换操作。第二,数据帧转发或过滤。将从输入介质上收到的数据帧转发至相应的输出介质或滤除。

在网桥中使用的路由选择技术就像网络层使用的那样,每个网桥中存储一张固定的路由表,选取的原则可以是某种既定的最短通路算法,还有两种路由策略:IEEE 802.1 发布的标准是基于生成树算法,可实现透明网桥;伴随 IEEE 802.5 标准的是源路由网桥规范。

如果将一个两端口的网桥改为多端口的设备,就是我们所说的交换机。交换机作为网桥的一种特殊形式,可以完成网桥能够完成的功能,并进行了改进,如端口数量的增多,有效增加了可用带宽,减小了碰撞域。

2. 交换机的地址学习及转发过程

交换机工作在 OSI 参考模型的第二层,它在运行过程中不断收集和建立自己的 MAC 地址表,并且定时刷新。如图 6.2 所示为网桥或交换机收到帧时的地址学习及转发过程。交换机的交换地址表中,一条表项主要由一个主机 MAC 地址和该地址所位于的交换机端口号组成。交换机仅根据 MAC 地址进行帧的选路和转发,当一个完整正确的以太网帧从一个交换机端口上被接收以后,交换机将在交换表查找收到 MAC 地址对应的端口号,然后再根据查找结果进行转发。

图 6.2 中的转发数据库,存放的是交换地址表。交换机采用动态自学习的方法,即当交换机收到一个数据帧以后,将数据帧的源地址和输入端口记录在交换地址表中。当一个数据帧从交换机某个特定端口 X 到达,交换机根据这个数据帧的信息可以得出:从端口 X 可以到达帧源地址域所指定的 PC 机,因此,交换机便将该 MAC 地址更新到转发数据库。为

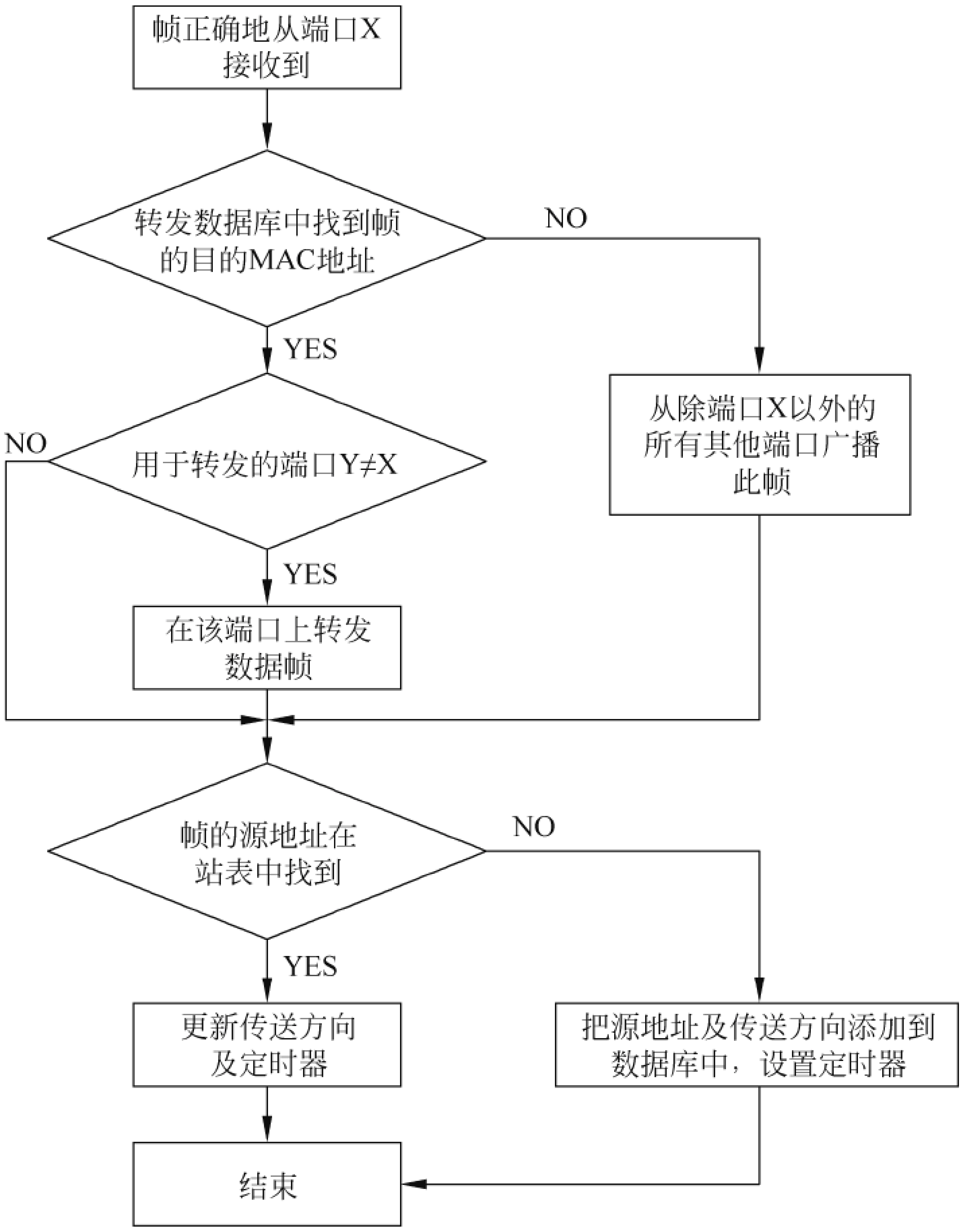


图 6.2 交换机的地址学习及转发过程

允许网络拓扑结构发生变化，转发数据库的每一项都配有寿命定时器，当一个新项加到数据库时，就启动定时器，定时器的默认值是 30s。如果定时器时间到，交换机就从转发数据库中搜索新项，判断新项中帧的源地址是否已存在：如果存在，转发数据库中项的内容被更新，并重新设置定时器值，如果转发数据库中不存在这样的项，将在转发数据库中添加一新项，该新项中的地址为收到数据帧的源 MAC 地址，端口号为收到数据帧的端口，定时器值被设置成初值。

6.1.2 数据帧转发与网段划分

1. 数据帧转发

交换机有 3 种数据包转发模式，即直通传送、存储-转发和改进型直通传送，如图 6.3 所示。

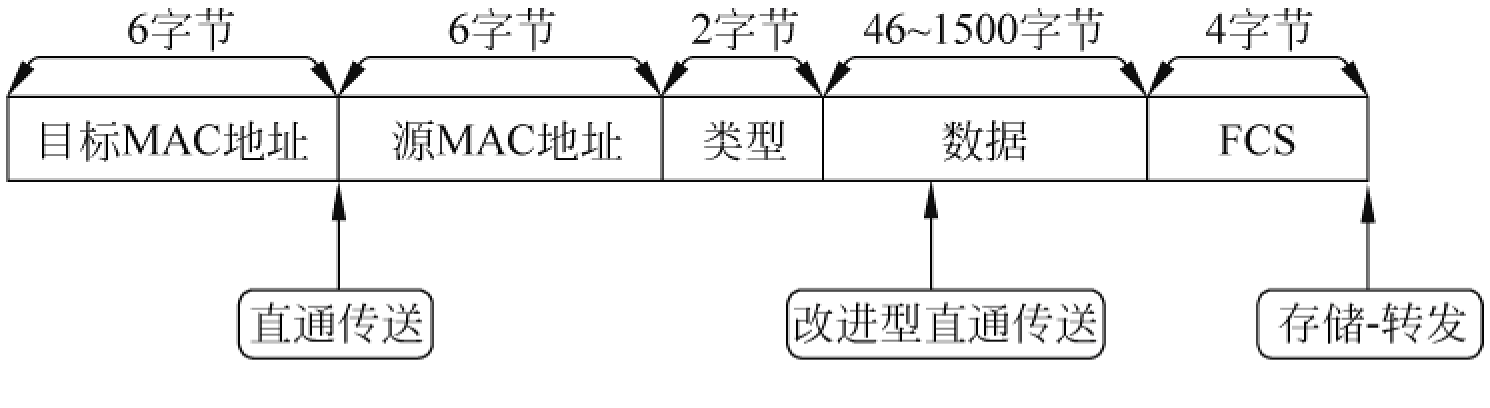


图 6.3 MAC 帧转发模式

(1) 直通传送。如图 6.3 所示,直通传送是指只要网桥收到数据帧的目的地址 MAC 字段,就立即将数据帧转发到相应的端口。因此,对数据帧的延迟很小,加大了数据包吞吐率。缺点是无法有效地检查出坏帧,目前大部分交换机都提供了直通传送的功能。

(2) 存储-转发。如图 6.3 所示,传统的网桥都用这种模式,网桥首先要将整个数据帧完全接收并存储下来,然后,根据数据帧的最后一个字段(帧校验序列)进行数据校验。如果校验正确再转发,否则丢弃收到的数据帧。采用这种方法的优点是可以有效地检查出来坏的数据帧,不足是转发速率慢。

(3) 改进型直通传送。如图 6.3 所示,改进型介于直通传送和存储-转发之间,它是等到正确收到数据帧的前 64 字节后开始进行转发。这样可以过滤收到的一些长度小于 64 字节的碎片帧。因此,这种方式也称为无碎片帧转发模式。

交换机转发数据帧时,遵循以下规则。

(1) 如果数据帧是单播帧(unicast),如目的地址在 MAC 地址表中存在,则按照目的地址在地址表中的表项所指的输出端口,将帧转发到相应的端口上(单播 MAC 地址在地址表中只能指向一个输出端口);如目的地址在 MAC 地址表中不存在,则在广播域的所有端口上广播该帧。

(2) 如果数据帧是多播帧(multicast),如目的地址在 MAC 地址表中存在,则按照目的地址在地址表中的表项所指的输出端口,将帧转发到相应的端口上(多播 MAC 地址在地址表中可以指向一个或一组输出端口);如目的地址在 MAC 地址表中不存在,则在广播域的所有端口上广播该帧。

(3) 如果数据帧是广播帧(broadcast),是指目的 MAC 地址为 ff-ff-ff-ff-ff-ff 的帧,要在广播域的所有端口上广播该帧。

(4) 如果数据帧的目的地址在交换机的地址表中,那么就根据地址表转发到相应的端口。

(5) 如果数据帧的目的地址与数据帧的源地址在一个端口上,它就会丢弃这个数据帧。

【例 1】 根据图 6.4 说明交换机的转发过程。

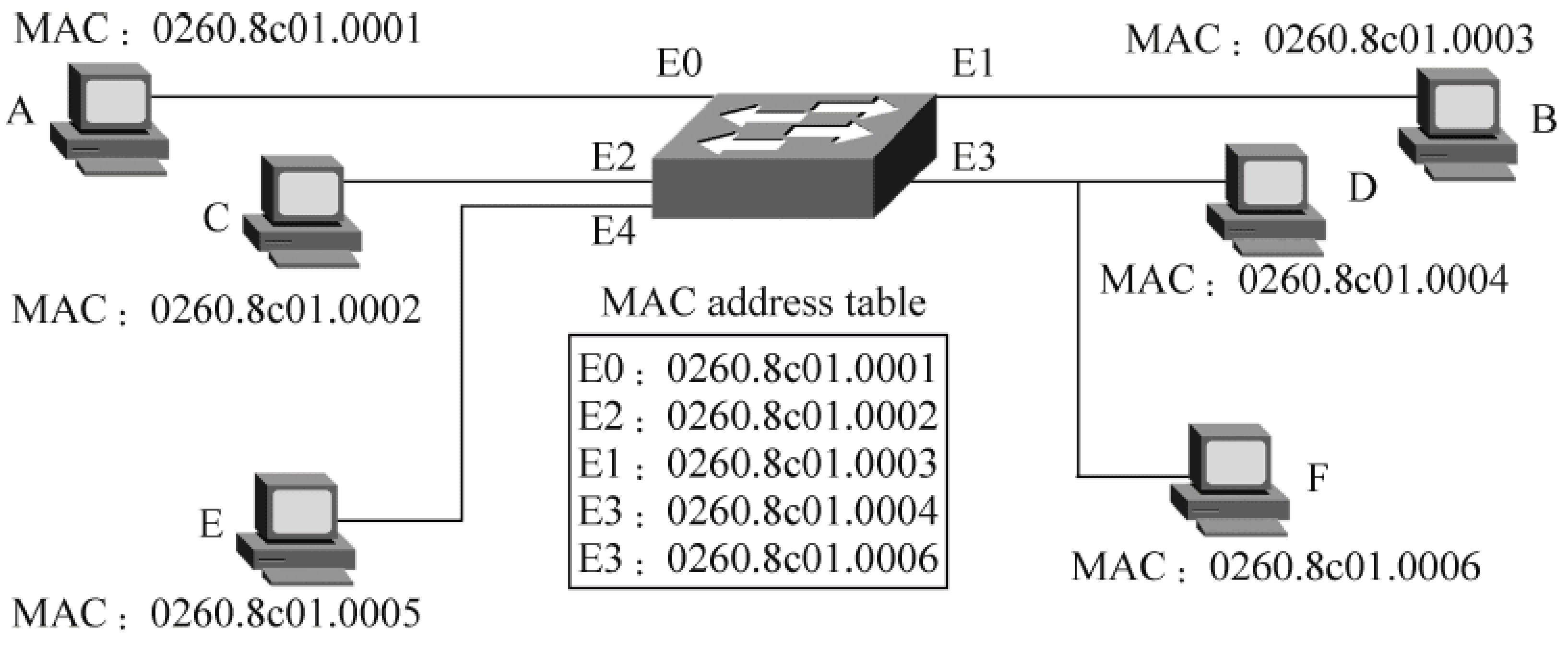


图 6.4 交换机的转发过程

答:当主机 D 发送广播帧时,交换机从 E3 端口接收到目的地址为 ffff.ffff.ffff 的数据帧,则向 E0、E1、E2 和 E4 端口转发该数据帧。

当主机 D 与主机 E 通信时,交换机从 E3 端口接收到目的地址为 0260.8c01.0005 的数据帧,查找地址表后发现 0260.8c01.0005 并不在表中,因此交换机仍然向 E0、E1、E2 和 E4 端口转发该数据帧。

当主机 D 与主机 F 通信时，交换机从 E3 端口接收到目的地址为 0260.8c01.0006 的数据帧，查找地址表后发现 0260.8c01.0006 也位于 E3 端口，即与源地址处于同一个端口，所以交换机不会转发该数据帧，而是直接丢弃。

当主机 D 与主机 A 通信时，交换机从 E3 端口接收到目的地址为 0260.8c01.0001 的数据帧，查找地址表后发现 0260.8c01.0001 位于 E0 端口，所以交换机将数据帧转发至 E0 端口，这样主机 A 即可收到该数据帧。

如果在主机 D 与主机 A 通信的同时，主机 B 也正在向主机 C 发送数据，交换机同样会把主机 B 发送的数据帧转发到连接主机 C 的 E2 端口。这时 E1 和 E2 之间，以及 E3 和 E0 之间，通过交换机内部的硬件交换电路，建立了两条链路，这两条链路上的数据通信互不影响，因此网络亦不会产生冲突。所以，主机 D 和主机 A 之间的通信独享一条链路，主机 C 和主机 B 之间也独享一条链路。而这样的链路仅在通信双方有需求时才会建立，一旦数据传输完毕，相应的链路也随之拆除。这就是交换机主要的特点。

我们发现，当交换机收到一个不认识的数据包时，也就是说如果目的 MAC 地址不在 MAC 地址表中，交换机便会把该包从所有端口“扩散”出去，如同收到一个广播包一样，这也是传统局域网交换机的弱点：不能有效地解决广播及安全性控制等问题。因此，就产生了交换机上的 VLAN(虚拟局域网)技术。

2. 网段(或 VLAN)划分

二层交换机可以将原以太网上的一个冲突域划分成多个冲突域。但是，交换机的所有端口仍处于一个广播域，当一台主机发送广播数据包时，所有主机都要接收并加以处理。当连接在交换机上的主机增加时，广播数据包将会大大降低网络的吞吐率。

特别是单个交换机连接了企业网络的不同部门时，一个部门的广播通信量将会传播到整个广播域，尽管一个部门大部分时候并不关心其他部门的广播数据。

可以采用路由器来隔离广播通信量，但路由器每端口的成本很高，不太可能为每个部门提供一个局域网接口。同时，因为路由器工作在网络层，这将会增加数据包的处理延迟。

为了在不改变设备和不增加传输延迟的情况下增加广播域的数目，可以采用虚拟局域网技术。虚拟局域网技术利用交换机把网络划分成若干个 VLAN，同时将广播通信量限制在每个 VLAN 内部，从而增加了广播域的数目，减少了广播对网络的影响。

如图 6.5 所示，由网络管理人员手工将交换机第 1～8 个端口设为 VLAN10 的成员，将第 9～16 个端口设为 VLAN20 的成员，将第 17～24 个端口设为 VLAN30 的成员。

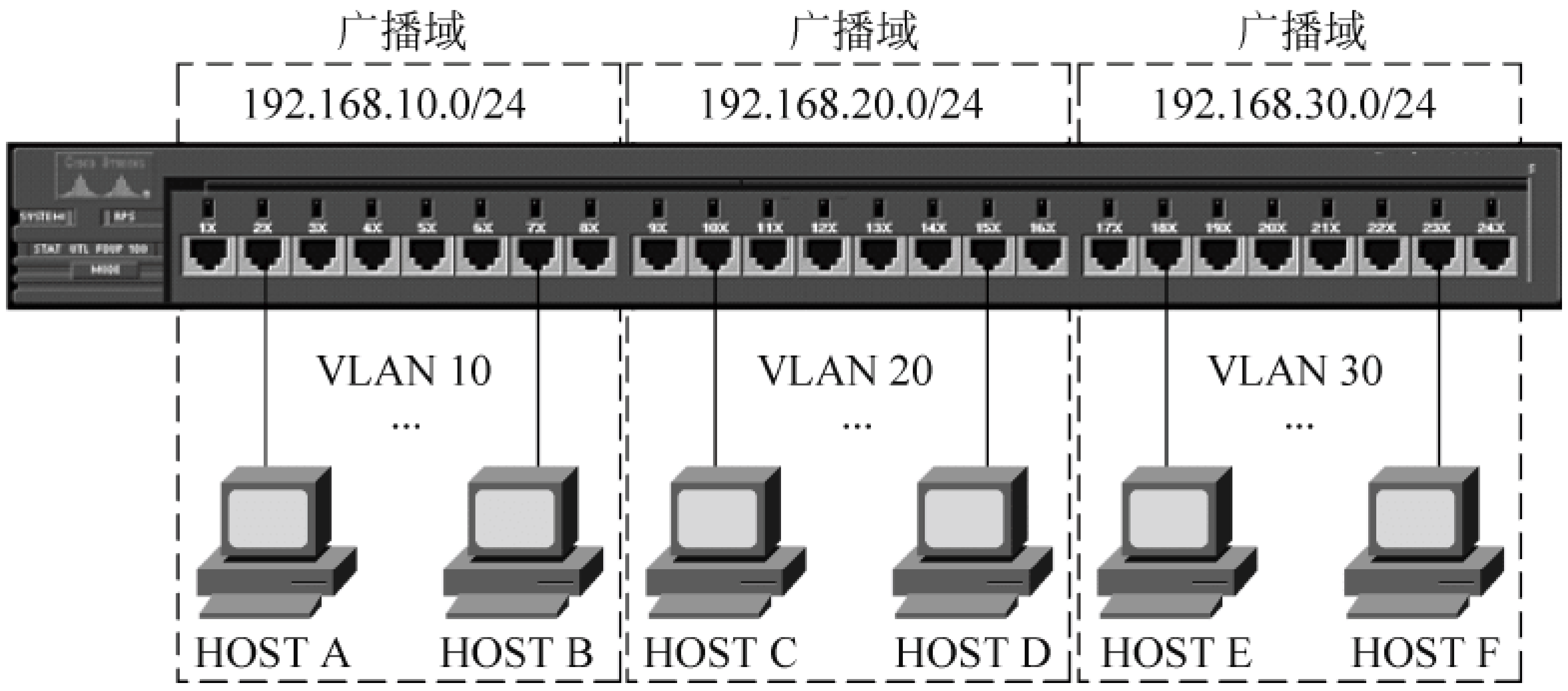


图 6.5　交换机网段划分

当交换机在 VLAN10 所属的端口收到广播数据帧后，它并不向 VLAN20、VLAN30 等其他 VLAN 所属的端口转发此广播数据帧。实际上，交换机在各 VLAN 之间不传输任何用户数据。如果各 VLAN 之间需要通信，必须借助外接的路由器或交换机内置的路由模块。

交换机的主要功能可概括为：交换机通过学习，可以得到源 MAC 地址；通过广播查询，可以找到目标 MAC 地址；通过过滤，可以使连接在一个端口的设备间的通信不会经过交换机转发；通过转发，会将从一个端口进入交换机的数据根据目标地址从另一个端口转发出去。

6.1.3　交换机互连

交换机主要功能包括物理编址、网络拓扑结构、错误校验、帧序列以及流量控制等。它使网络各站点之间可独享带宽，消除了无谓的冲突检测和出错重发，提高了传输效率，而且是点对点传送用户数据，其他节点是不可见的。

交换机有不同厂家的多种产品，如图 6.6 所示为 Cisco 交换机的前面板，型号为 2950-24，它的前面板除了有供接入主机的 24 个 RJ-45 端口外，还有很多状态指示灯，如端口指示灯(PORT)、系统指示灯(SYST)、冗余电源指示灯(RPS)、状态指示灯(STRT)、双工指示灯(DUPLX)、速率指示灯(SPEED)和利用率指示灯(UTIL)等。

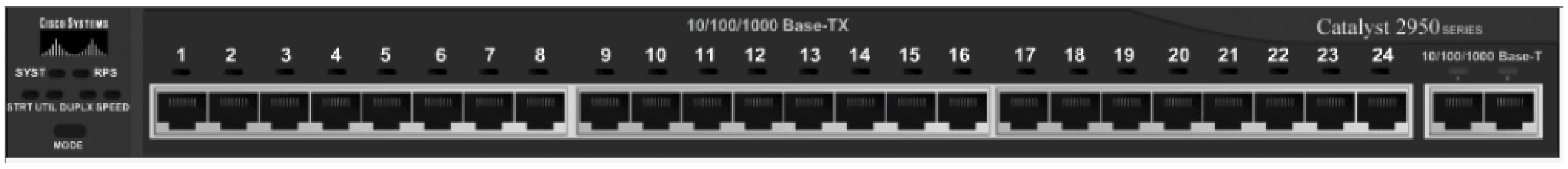


图 6.6　交换机前面板

交换机一般除了固有的配置以外，还可以根据需要增加配置多种模块，如：NM-1FE-FX 模块提供一个快速以太网接口，介质是光纤；NM-1FE-TX 模块提供一个支持铜介质的快速以太网接口。它们都适合于组建远距离局域网应用，快速以太网模块支持多种特性和标准。

其中，NM：网络模块(Network Module)。

1FE：1 个 FE(FastEthernet)接口，速率为 1000Mbps。

TX：是普通的电接口(铜介质)，即 RJ45 网口。TX 的版本支持虚拟局域网扩展。

FX：是光纤接口，如：ST 接口(10Base-F)，SC 接口(100Base-FX)等。

通常将若干交换机互连起来以增加端口密度，有以下两种不同的交换机互连方法。

1. 级连方法

采用级连方式连接交换机时，只需要用双绞线直接将两台交换机的某两个普通端口连接起来就可以了。级连后，交换机之间的数据交换线路带宽就是级连端口的带宽。当把同类型的交换机通过普通端口互连时，必须使用交叉线序的双绞线(Crossover UTP)。

有的厂商一般将交换机上的最后一个端口设为级连端口(MDI，Medium-Dependent Interface，介质有关的接口)，并标为 MDI/MDI-X，同时提供一个按钮进行 MDI 模式/普通端口模式之间的转换。当为 MDI 模式时，可用直通线序双绞线直接将两台交换机相连。还有的厂商将 MDI/MDI-X 接口设置为可以自动协商，并可智能调整接口内部线序。

2. 堆叠方法

由于级连链路带宽有限,有的厂商就提供了另外一种互连方式:交换机的堆叠扩展。使用堆叠技术进行扩展的交换机,必须有单独的专用堆叠接口和专用的堆叠电缆,堆叠接口提高了交换机间链路带宽,速率一般都高于普通接口。

6.2 多层交换机

由于以太网交换技术不能有效解决广播风暴、网络互连和安全性控制等问题,又推出了三层以上的多层交换机,将广播和本地流量限制在一定的范围内。以下主要介绍三层交换机。

6.2.1 交换技术

1. 交换技术概述

通常将以太网交换机构成的局域网称为交换式以太网,为了减小广播域,可以采取划分虚拟局域网(VLAN)的方式来实现,即每个以太网交换机划分为一个 VLAN,VLAN 间的互通则需要由路由器来完成,如图 6.7 所示。但传统路由器配置复杂,受转发速度的限制,也易成为网络的瓶颈,所以能代替路由器的、具有路由功能的三层交换机就呼之欲出。

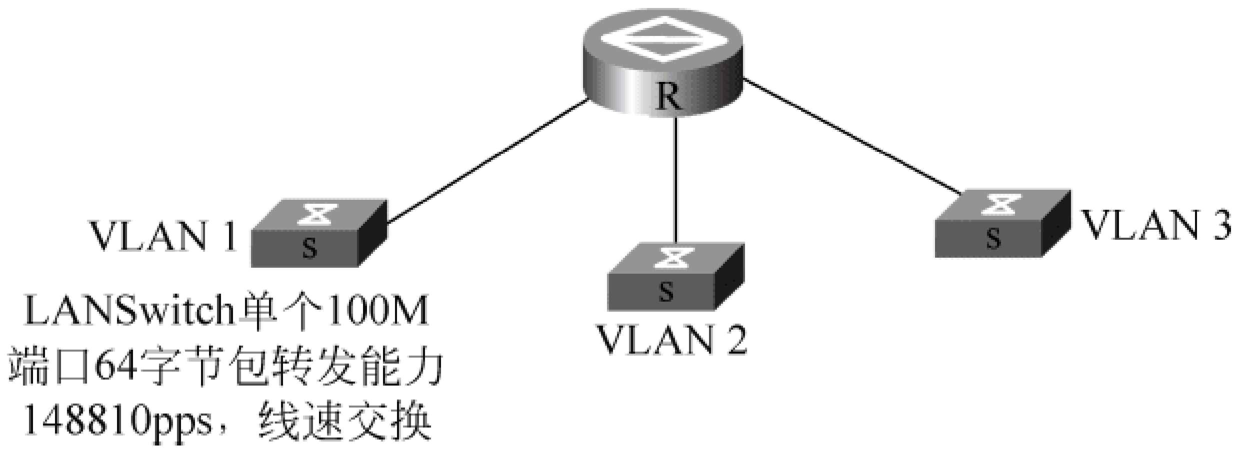


图 6.7 路由器连接的以太网交换机

三层交换机与传统路由器的比较:三层交换机的实质就是通过 ASIC 硬件进行包转发,它与传统的路由器具有相同的功能,而传统路由器是采用通用处理器芯片,通过运行软件来实现的路由,它们对于数据包的处理过程是相同的,都需要根据目标 IP 转发报文,改变报文中的 MAC 地址,递减 IP 头部中的生存时间值,进行数据校验等。

三层交换机也适用于静态路由、动态路由来创建和维护路由表。但是,交换机为了实现高速转发数据包,通常采用"路由一次,交换多次"的工作方式。即去往同一网络的数据包只在第一次被路由器处理时才查找路由表,其后续的去往同一目标的数据包将被直接传送到输出接口,以提高转发速率。

三层交换机共分为接口、交换和路由 3 个部分。接口部分包含了所有重要的局域网接口,如 10/100/1000Mbps 以太网、FDDI 和 ATM 等;交换部分集成了多种局域网接口,并辅之以策略管理,同时还提供链路汇聚、VLAN 和标签机制;路由部分提供主要的局域网路由协议,包括 IP 等。

由于路由器生成树协议收敛速度较慢,当网络出现链路故障时,选路灵活性也会受到限制,所以在跨越 VLAN 时采用三层交换机来取代路由器就成为人们的必然选择。

2. 三层交换技术

三层交换也称为 IP 交换、高速路由技术等，三层交换主要涉及如下技术。

(1) Ipsilon IP 交换：提倡识别数据包流，尽量在第二层进行交换，以绕过路由器，改善网络性能。该技术适用于机构内部的局域网和校园网。

(2) Cisco 标签交换：给数据包贴上标签，此标签在交换节点读出，判断包传送路径。该技术适用于大型网络和 Internet。

(3) 3Com Fast IP：侧重数据策略管理、优先原则和服务质量。Fast IP 协议保证实时音频或视频数据流能得到所需的带宽，并支持其他协议。

(4) IBM ARIS：与 Cisco 的标签交换技术相似，包上附上标记，借以穿越交换网。ARIS 一般用于 ATM 网，也可扩展到其他交换技术。后来 IETF 在前面基础上又推出 MPLS(多协议标记交换)，为各种网络层协议和数据链路层协议提供了一种有效的多协议解决方案，以满足网络服务质量要求并提供流量工程支持。

(5) MPOA：是 ATM 论坛提出的一种规范。经源客户机请求，路由服务器执行路由计算后给出最佳传输路径。然后，建立一条交换虚电路，即可越过子网边界，不用再做路由选择。

6.2.2　交换原理

第三层交换可直接利用动态建立的 MAC 地址来通信，具有多路广播和虚拟网间基于 IP 等协议的路由功能。

1. 基本交换原理

三层交换基本原理如图 6.8 所示。图中，假设主机 A 和主机 B 要通信，A 首先向交换机发送一个 ARP 请求包，寻找自己默认路由的目的 MAC 地址，然后将数据发送到交换机，若 A 和 B 在同一个子网中，直接通过交换机的二层将数据包转发出去，不再经过三层。如果 A 和 B 不在同一个子网中，需要将数据包转发到三层，在路由表中寻找匹配的条目。若在表中找到相关条目，则通过 MAC 地址，直接在二层建立连接，使芯片处理数据通过二层转发数据包，既节省时间，又提高效率；若在表中无法找到相关项，需三层交换机将目的 IP 地址和路由表逐项对比，并发送 ARP 广播数据包到目的主机 B，得到该主机的 MAC 地址，然后在二层转发数据。三层交换机主要用于子网的连接，简单概括如下：

(1) 第一种情况是 A、B 在同一子网，这时候，三层交换机作为以太网交换机使用。

(2) 第二种情况是 A、B 不在同一个子网，但 B 直连在三层交换机的端口上，直接找到目地 MAC 地址进行交换。

(3) 第三种情况是三层交换机连接一个以太网交换机，B 就在这个以太网交换机子网中，三层交换机需要通过路由表发送 ARP，找到目的 MAC 地址，然后在二层转发数据。

三层交换机将二层交换机和路由器的功能集成在一起，在三层交换机中分别体现为二层交换引擎和三层路由引擎两个部分功能。

(1) 二层交换引擎：实现同一网段内的快速二层转发，是用硬件支持多个 VLAN 的二层转发。

(2) 三层路由引擎：实现跨网段的三层路由转发，是使用硬件 ASIC 技术实现高速的 IP 转发。

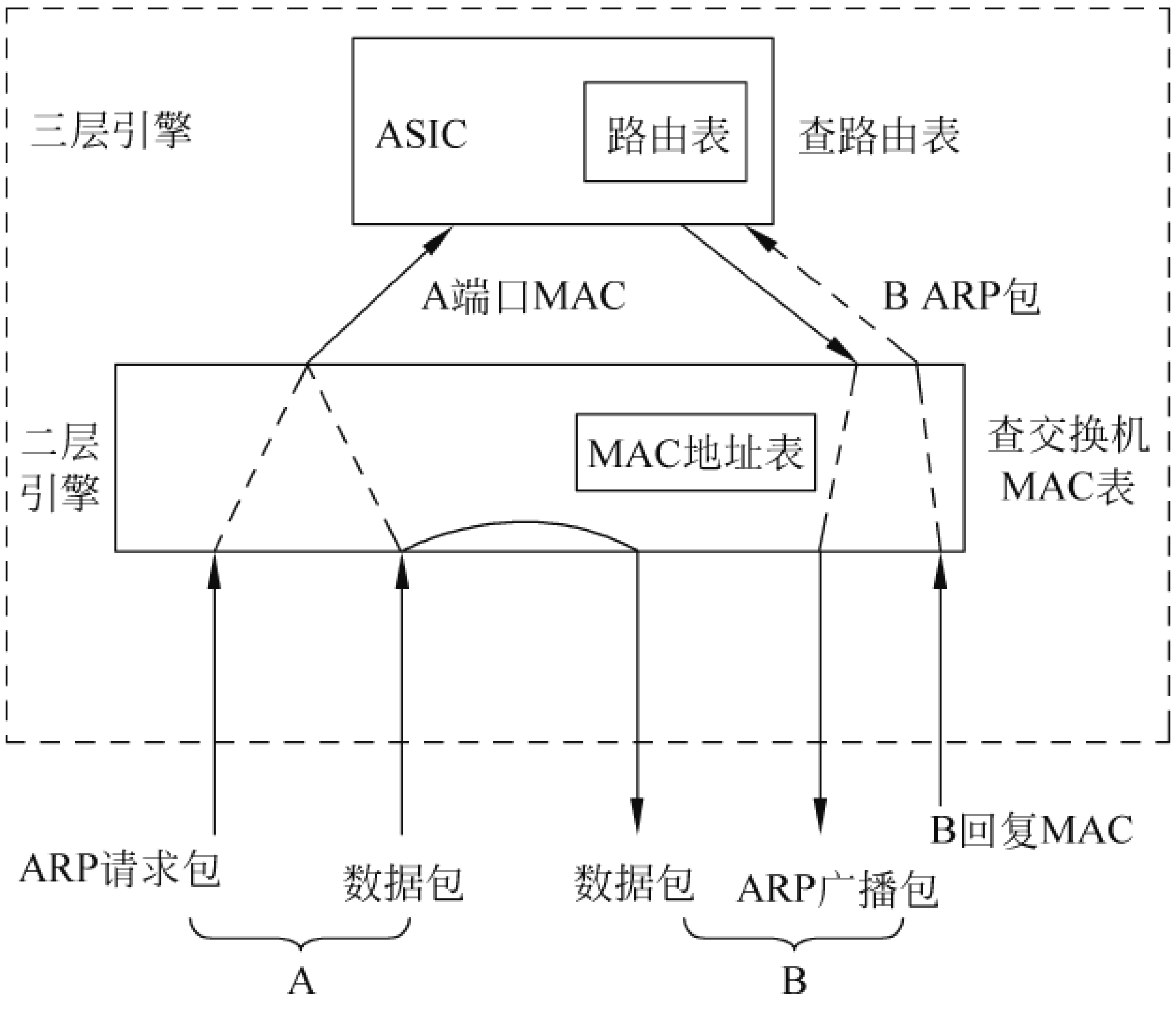


图 6.8 三层交换基本原理

在 IP 网络中，每个 VLAN 对应一个 IP 网段，三层交换机中的三层转发引擎在各个网段(VLAN)间转发报文，实现 VLAN 之间的互通，因此三层交换机的路由功能通常叫做 VLAN 间路由(Inter-VLAN Routing)。

图 6.9 给出了三层交换机转发流程，数据包进入三层后要经过多层流分类、处理等流程。

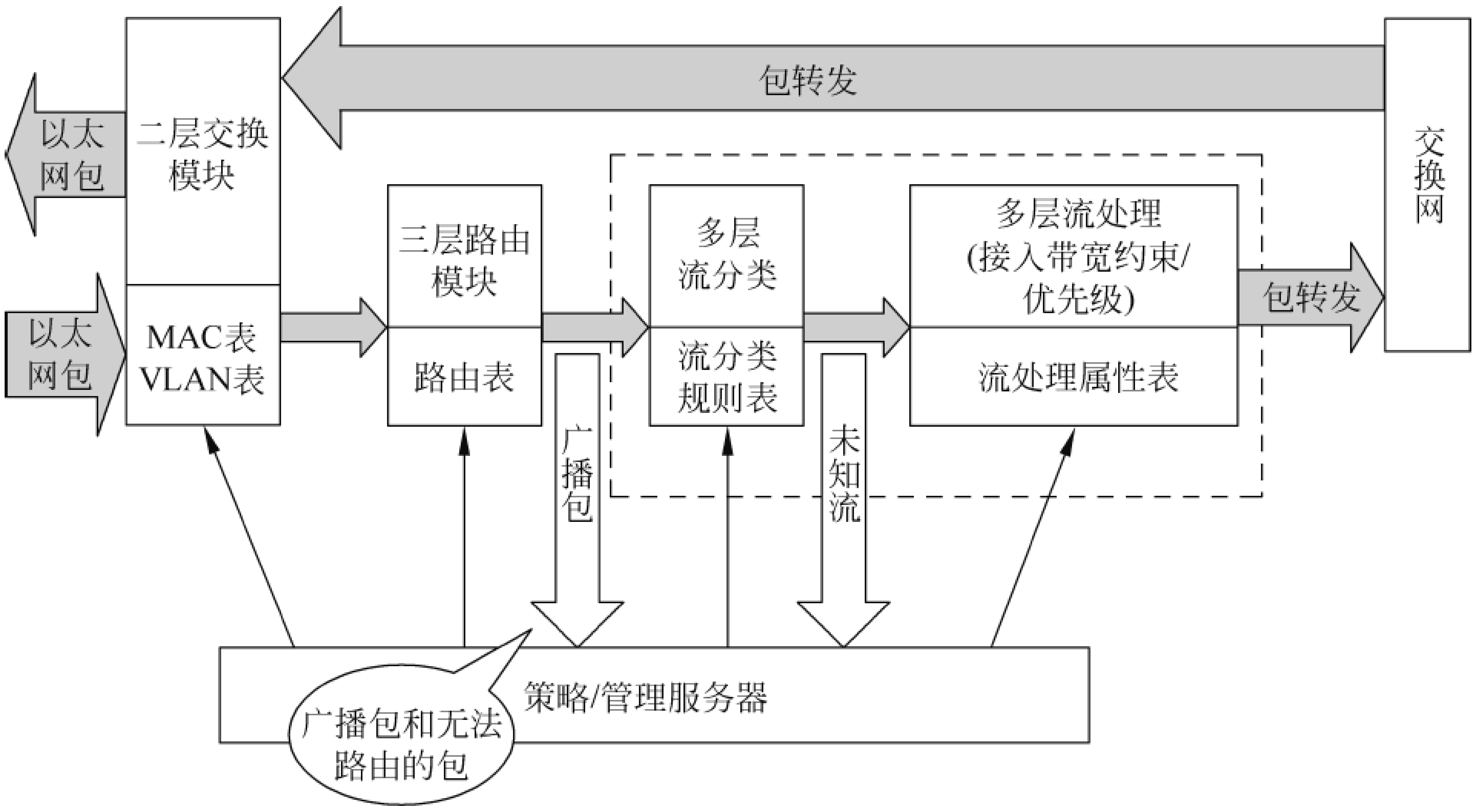


图 6.9 三层交换机转发流程

在二层上，VLAN 之间是隔离的。这一点与二层交换机中的交换引擎的功能一模一样。不同的 IP 网段之间的访问要跨越 VLAN，要使用三层转发引擎提供的 VLAN 间路由功能(相当于路由器)。

在使用二层交换机和路由器的组网中，每个需要与其他 IP 网段(VLAN)通信的 IP 网段(VLAN)都需要使用一个路由器接口做网关。三层转发引擎就相当于传统组网中的路由器功能，当需要与其他 VLAN 通信的时候也要为之在三层交换引擎上分配一个路由接口，

用来做 VLAN 的网关。这个路由接口是在三层转发引擎和二层转发引擎上的，是通过配置转发芯片来实现的，与路由器的接口不同，这个接口不是直观可见的。

在 VLAN 指定路由接口的操作，实际上就是为 VLAN 指定一个 IP 地址、子网掩码和 MAC 地址，MAC 地址是由设备制造过程中分配的，在配置过程中由交换机自动配置。

三层交换对 VLAN 之间的通信的解决方法为：在 VLAN 之间配置路由器，这样 VLAN 内部的流量仍然通过原来的 VLAN 内部的二层网络进行，从一个 VLAN 到另外一个 VLAN 的通信流量通过路由在三层上进行转发，转发到目的网络后，再通过二层交换网络把报文最终发送给目的主机。

由于路由器对以太网上的广播报文采取不转发的策略，因此中间配置的路由器仍然不会改变划分 VLAN 所达到的广播隔离的目的。ARP 是一种广播协议，主机通过它可以动态地发现对应于一个特殊 IP 网络层地址的 MAC 层地址。三层交换机会自动维护一张这样的 MAC-IP 地址对应表。三层交换过程如图 6.10 所示，IP 报头中的 IP 源地址与目的地址始终不变，但在第二层，数据帧的源 MAC 地址和目的 MAC 地址会根据 MAC-IP 地址对应表进行重新修改。

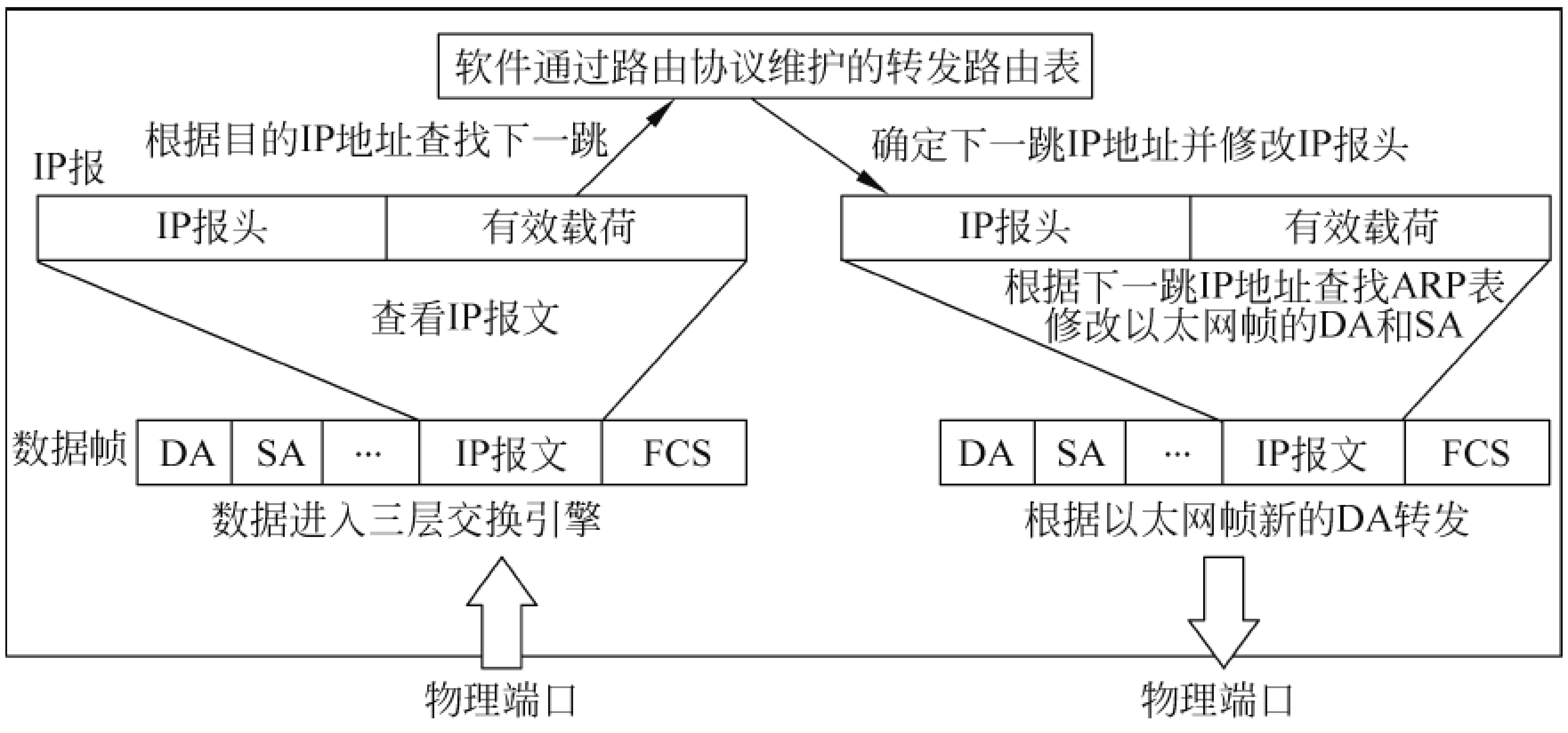


图 6.10　三层交换过程

【例 2】 假设两个的站点(源站点 A、目的站点 B)通过三层交换机连接，要求详细叙述其交换转发的通信过程。

答：

(1) 源站 A 在开始发送时，已知目的站的 IP 地址，但尚不知道在局域网上发送所需要的 MAC 地址。首先要发送地址解析(ARP)来确定目的站的 MAC 地址。

(2) 若目的站 B 与源站 A 在同一子网内，源站 A 广播一个 ARP 请求，目的站 B 返回其 MAC 地址，A 站得到目的站 B 的 MAC 地址后将这一地址缓存存放在 ARP 表中，并用此 MAC 地址封装包后转发数据。三层以太网交换机的第二层交换模块(可以理解为同一子网中的二层交换机)根据源站 A 发送的以太网帧中的目的 MAC 地址查找 MAC 地址表确定将数据包发向目的端口。

(3) 若目的站 B 与源站 A 不在同一子网内，站 A 需要与目的站 B 联系，站 A 要向"默认路径(配置的网关地址)"发出 ARP。当站 A 对"默认路径"的 IP 地址广播出一个 ARP 请求时，交换机返回相应路由接口(即源站 A 的"默认路径")的 MAC 地址给源站。

(4) 若第三层交换模块在以往的通信过程中已得到目的站 B 的 MAC 地址，则发向目的站 B 的数据包可直接封装此 MAC 地址。否则，提取出输入帧的 IP 包去查路由表，根据路由表中的路由信息向目的站 B 网段广播一个 ARP 请求，目的站 B 得到此 ARP 请求后，向第三层交换模块回复其 MAC 地址，以后，当再进行站点 A 与站点 B 之间的数据包转发时，将用最终的目的站点 B 的 IP 地址为索引查找底层硬件转发表，得到出端口与对应的 MAC 地址，并用查到 MAC 地址封装包，从查到的出端口将数据转发出去，从而数据转发过程全部交给第二层交换处理，因此信息得到高速交换。

2. 报文交换

传统三层技术就是对每个报文进行处理，并基于第三层 IP 地址转发报文，这一方法称为报文到报文的三层交换技术，如图 6.11 所示。

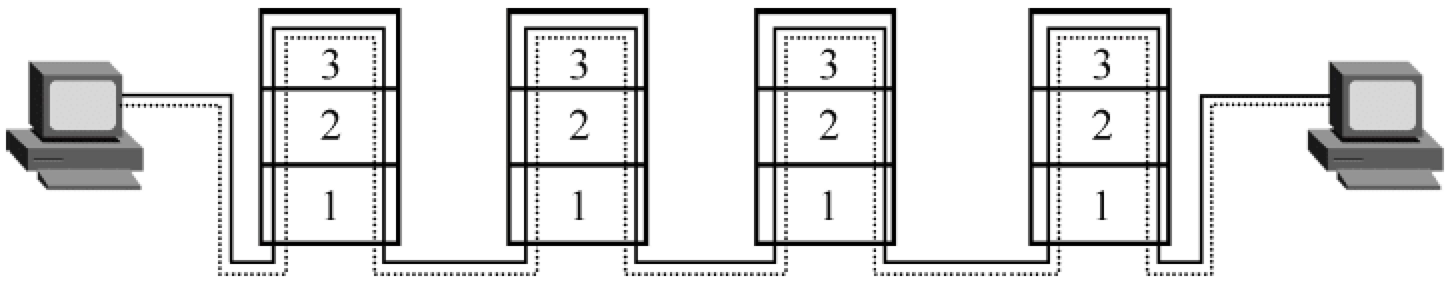
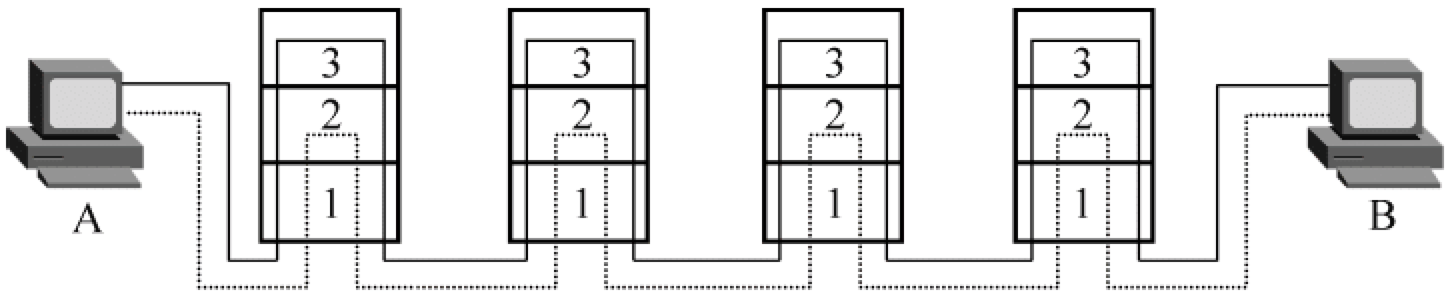


图 6.11 报文到报文的三层交换技术

在报文到报文的交换中，报文进入交换机的第一层物理接口，然后到达第二层接口进行目的 MAC 地址检查，如果查表检查的结果是不能交换则进入到第三层。在第三层，报文通过路由计算、地址解析等处理，经过三层处理后，报文头被修改并被传回第二层，第二层确定合适的输出端口后，报文通过第一层传送到物理介质上。对于后续的每一个报文的转发，都要经过这样的一个过程。

3. 流交换

不在第三层处理所有报文的方法称为流交换(FS)。报文到报文的交换方式与流交换方式的区别为：如果每一个报文都要经历第三层处理，并且业务流转发是基于第三层 IP 地址的，这种交换方式就是报文到报文交换方式；如果只是第一个报文经过第三层处理，其他后续报文只进行第二层转发，这种交换方式就是流交换方式。基于流交换的三层交换如图 6.12 所示。



图中：“—”表示第一个报文“……”表示后续报文(流)

图 6.12 基于流交换的三层交换

在流交换中，第一个报文被分析以确定其是否表示了一个“流”或者一组具有相同源地址和目的地址的报文，如果第一个报文具有正确的特征，则该标识流中的后续报文将拥有相同的优先权，同一流中的后续报文被交换到基于第二层的目的地址，流交换节省了检查每一个报文要花费的处理时间。现在三层交换机为了实现高速交换，基本上都采用流交换的方式。

4. 背板带宽

一台交换机的背板带宽体现了它的处理能力，应从以下两个方面判断交换机提供背板带宽的可用性。

(1) 如果满足(全部端口容量×端口数量)×2≤背板带宽，可实现全双工无阻塞交换，说明交换机具有发挥最大数据交换性能的条件。

(2) 如果满配置吞吐量(Mpps)=满配置 GE 端口数×1.488Mpps，这里是以千兆口交换机为例，其中 1 个千兆端口(GE)在包长为 64 字节时的理论吞吐量为 1.488Mpps。

已知以太网传输最小包长是 64 字节，包转发线速的衡量标准是以单位时间内发送 64 字节的数据包的个数作为计算基准的，因此对一个全双工线速的千兆以太网端口在转发 64 字节包时的包转发率为：

$$1000\text{Mbps}/((64\text{ 字节}+8\text{ 字节}+12\text{ 字节})\times 8\text{ 位})=1.488\text{Mpps}$$

以上计算中，当以太网帧为 64 字节时，需考虑 8 字节的前导符，它的的作用是告诉监听的接收设备后面有数据要到来。当然还要考虑帧间隙的固定开销，在以太网标准中规定最小间隔为 12 字节。因此，常用以太网端口的包转发率：万兆以太网为 14.88Mpps；千兆以太网为 1.488Mpps；百兆以太网为 0.1488Mpps。

6.3 交换机配置

使用交换机配置向导只能完成一些基本的初始配置。对于更为详细的参数、选项设置，包括交换机名称、加密使能密码、虚拟终端密码、控制台密码等，要通过手工配置的方式来完成。本节将介绍 CISCO 基本命令，并利用交换机进行有关 VLAN 的配置。

6.3.1 常用配置命令

1. 交换机基本状态和模式

```
switch                              !ROM 状态
switch>                             !用户模式
switch#                             !特权模式
switch(config)#                     !全局配置模式
switch(config-if)#                  !接口状态
Switch(config-line)#                !控制线状态
```

2. 交换机基本常用配置命令

```
switch>enable                       !进入特权模式
switch #configure terminal          !进入全局模式
switch(config)#hostname Hostname    !设置交换机的主机名：Hostname
switch(config)#interface serial 1/1 !进入接口：s1/1
switch(config)#enable secret xxx    !设置特权加密口令：xxx
switch(config)#enable password yyy  !设置特权非密口令：yyy
switch(config)#line console 0       !进入控制台口(或称为控制线)
switch(config-line)#login           !允许登录
switch(config-line)#password xxx    !设置登录口令：xxx
switch(config)#line vty 0 4         !远程登录虚拟端口 5 个，最多为 0 到 15 一共 16 个终端
```

```
switch(config)# interface vlan 100                        !设置接口对应的 VLAN: 100
switch(config-if)# ip helper-address 20.2.2.2  !在 VLAN 接口状态下,设置 VLAN 动态配置地址
switch(config)# interface FastEthernet 0/1                !进入接口: f0/1
switch(config-if)# cdp enable                             !CDP 使能,启用和浏览 CDP 信息
switch(config-if)# no cdp enable                          !关闭 CDP 使能
switch(config-if)# description description-string         !接口描述: description-string
switch(config-if)# switchport mode access                 !将接口设置为 access 模式
switch(config-if)# switchport mode trunk                  !将接口设置为 trunk 模式
switch(config-if)# speed auto                 !速率设为自动,可改变速率为{10|100|auto}
switch(config-if)# duplex full                !以太网链路设为全双工模式,可以选{auto|full|half}
```

例如,对远程登录用户设置如下。

```
switch(config)# enable secret 1234                        !设置进入特权模式进的密码
switch(config)# line vty 0 5                  !设置远程登录虚拟端口,这时可以同时打开 6 个会话
switch(config-line)# password 5678                        !设置登录虚拟端口密码为 123456
switch(config-line)# login                                !远程登录
```

3. 交换机常用显示命令

交换机可以使用 MS-DOS 的 ping、tracert 测试交换机配置是否正确。除此之外,也可以使用一些 IOS 命令对交换机配置进行检查。

```
switch# show cdp interface FastEthernet 0/1               !查看邻接设备的 CDP 通告信息
switch# show running-config                               !显示交换机运行配置文件内容
switch# show startup-config                               !该命令显示交换机启动配置文件内容
switch# show interface vlan 100                           !显示 VLAN100 是否激活、交换机 MAC 地址
switch# show vlan                                         !查看 VLAN 配置信息
switch# show mac address-table                            !显示 MAC,即桥接表内容
switch# write                                             !保存配置信息
switch# dir flash                                         !查看交换机闪存
switch# show vtp                                          !查看 vtp 配置信息
switch# show interface                                    !显示接口是否激活、接口 MAC 地址等参数
switch# show interface f0/0                               !查看指定接口 f0/0 信息
switch# show ip interface brief               !检查交换机的接口状态,brief 为接口类型+接口编号
switch# clear mac-address-table dynamic                   !清除动态学习的 MAC
switch# show cdp neighbors                                !查看该交换机有几个邻居
switch# clear cdp table                                   !清除 CDP 邻居消息
switch# copy running-config startup-config                !保存设置文件
switch# erase startup-config                              !清空启动设置文件
```

在具有三层功能的以太网交换机配置路由时,网络管理员可以通过静态路由配置方法将其设置,也可以运行路由协议实现动态配置。用 ip routing 命令可以打开和路由器几乎相同的路由功能,也可以用 no switchport/switchport 命令改变接口的属性。

【例 3】 交换机 Switch0 的网络拓扑图如图 6.13 所示,要求用 show 命令查看有关状态。

答:

(1) 执行 show ip interface brief 命令检查交换机的接口状态,状态信息如图 6.14 所示,其中 unassigned 表示未赋值,manual 表示手动配置。

(2) 执行 show mac-address-table 命令检查交换机 MAC 地址表如图 6.15 所示。

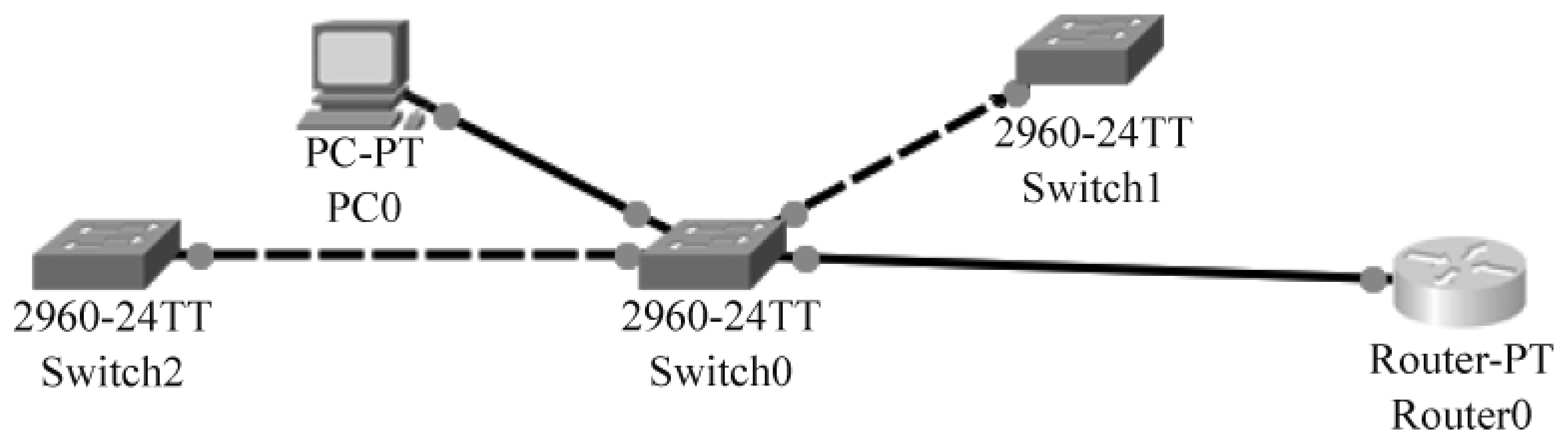


图 6.13　网络拓扑结构

```
Switch#show ip interface brie
Interface              IP-Address      OK? Method Status                Protocol

FastEthernet0/1        unassigned      YES manual up                    up

FastEthernet0/2        unassigned      YES manual up                    up

FastEthernet0/3        unassigned      YES manual up                    up

FastEthernet0/4        unassigned      YES manual up                    up

FastEthernet0/5        unassigned      YES manual down                  down
```

图 6.14　交换机的接口状态

```
Switch#show mac-address-table
          Mac Address Table
-------------------------------------------

Vlan    Mac Address       Type        Ports
----    -----------       --------    -----

   1    0006.2a56.3101    DYNAMIC     Fa0/2
   1    00d0.58c0.8d01    DYNAMIC     Fa0/1
   1    00d0.d320.b6c4    DYNAMIC     Fa0/4
```

图 6.15　交换机 MAC 地址表

（3）执行 Switch＃show cdp neighbors 命令后，查看该交换机的邻居状况如图 6.16 所示，其中：

```
Switch#show cdp neighbors
Capability Codes: R - Router, T - Trans Bridge, B - Source Route Bridge
                  S - Switch, H - Host, I - IGMP, r - Repeater, P - Phone
Device ID     Local Intrfce   Holdtme    Capability   Platform    Port ID
Switch        Fas 0/1          129            S       2960        Fas 0/1
Switch        Fas 0/2          141            S       2960        Fas 0/1
Router        Fas 0/4          164            R       PT1000      Fas 0/0
```

图 6.16　交换机邻居信息

Device ID 为所连接的设备名称；

Local Intrfce 指本地设备接口；

Holdtme 是保持时间；

Platform 是所连接对端设备的型号；

Port ID 是连接对端设备的接口；

Capability 列出所连设备的具体代码，设备代码（Capability Codes）所代表的意思如下：

R—Router(路由器),T—Trans Bridge(网桥),B—Source Route Bridge(源路由),S—Switch(交换机),H—Host(主机),I—IGMP,r—Repeater(中继器),P—Phone(电话)。

6.3.2 配置实例

1. VLAN 间路由选择

要实现 VLAN 间的数据传递,可以使用外接路由器或内置路由模块的方法来实现。最简单的一种实现方法就是利用多个路由器接口实现 VLAN 间路由选择,如图 6.17 所示,对交换机和路由器几乎不需要额外的配置,在每个 VLAN 中选出一个接口,如 VLAN10 中的 fastethernet0/1、VLAN20 中的 fastethernet0/9、VLAN30 中的 fastethernet0/17 分别通过双绞线接入路由器接口 fastethernet0/0、fastethernet0/1、fastethernet0/2 即可。在交换机上不需要配置任何选项。在路由器端也不需要配置路由信息(静态或动态路由配置),因为路由器会为其直连接口 fastethernet0/0、fastethernet0/1、fastethernet0/2 产生直连接路由,而它们所对应的 IP 地址和子网掩码以及每个 VLAN 中每台主机的 IP 地址、子网掩码和默认网关还是需要配置的。本实例综合应用了路由器、交换机等设备对 VLAN 的组网,对其设置如下。

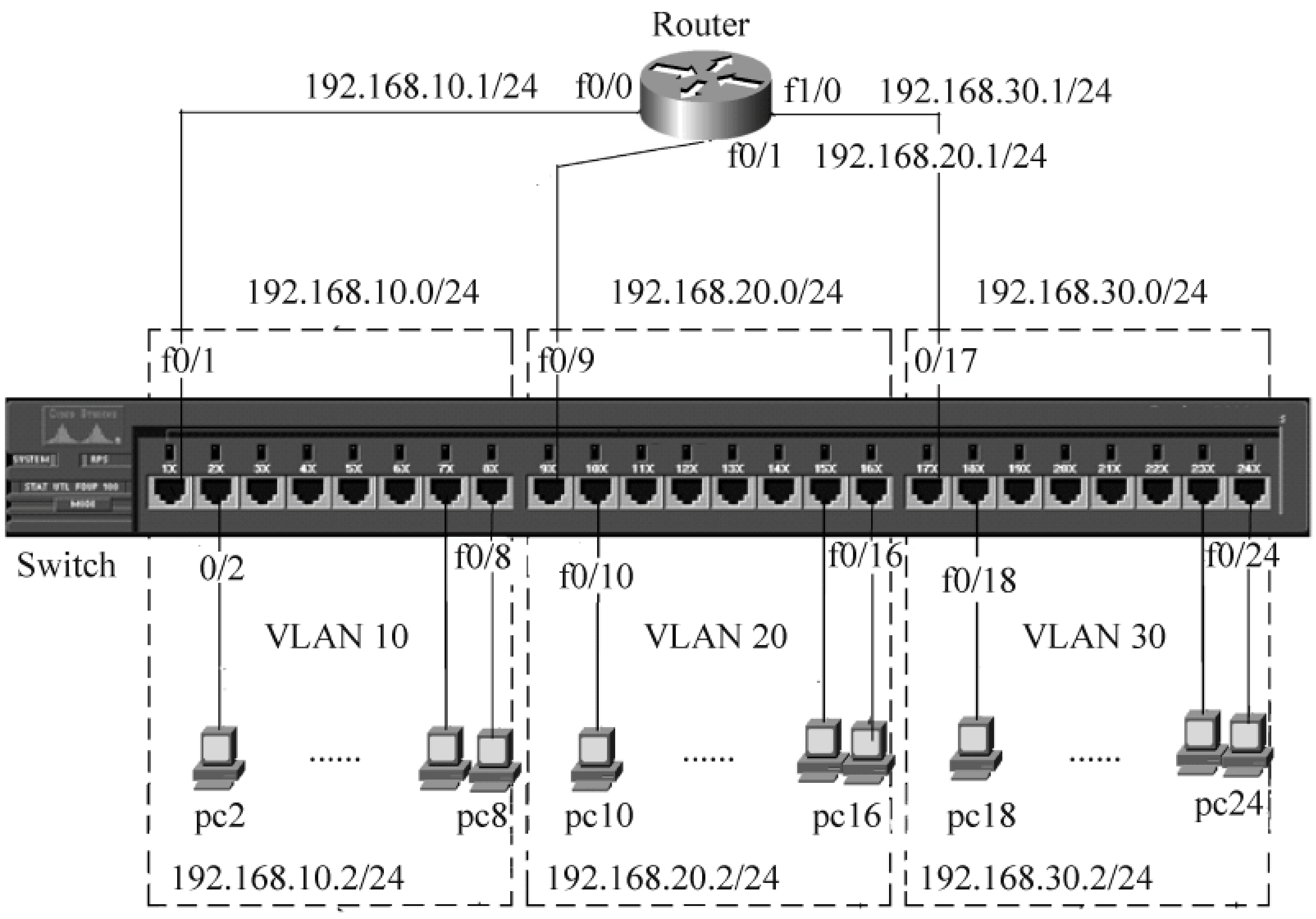


图 6.17 利用多个路由器接口实现 VLAN 间路由选择

1) 路由器配置

```
Router>enable
Router#config terminal
Router(config)#interface f0/0                 !进入并配置 fastethernet0/0 的 IP 地址
Router(config-if)#ip address 192.168.10.1 255.255.255.0
Router(config-if)#no shutdown
Router(config-if)#exit
Router(config)#interface f0/1                 !进入并配置 fastethernet0/1 的 IP 地址
Router(config-if)#ip address 192.168.20.1 255.255.255.0
```

```
Router(config-if)#no shutdown
Router(config-if)#exit
Router(config)#interface f1/0                    !进入并配置 fastethernet1/0 的 IP 地址
Router(config-if)#ip address 192.168.30.1 255.255.255.0
Router(config-if)#no shutdown
Router(config-if)#exit
```

2）交换机配置

```
Switch>enable
Switch#config terminal
Switch(config)#vlan 10                              !建立 VLAN 10
Switch(config-vlan)#vlan 20                         !建立 VLAN20
Switch(config-vlan)#vlan 30                         !建立 VLAN30
Switch(config-vlan)#interface range f0/1-8          !指定接口范围
Switch(config-if-range)#switchport mode access      !确定为访问模式
Switch(config-if-range)#switchport access vlan 10   !加入 VLAN10
Switch(config-if-range)#exit
Switch(config-if)#interface range f0/9-16           !指定接口范围
Switch(config-if-range)#switchport mode access      !确定为访问模式
Switch(config-if-range)#switchport access vlan 20   !加入 VLAN20
Switch(config-if-range)#exit
Switch(config-if)#interface range f0/17-24          !指定接口范围
Switch(config-if-range)#switchport mode access      !确定为访问模式
Switch(config-if-range)#switchport access vlan 30   !加入 VLAN30
Switch(config-if-range)#end
Switch#
```

3）PC 设置

对每个 PC 都要进行设置，根据 VLAN 地址，设置 PC 地址和掩码，网关地址就是接入路由器接口的 IP，如图 6.18 所示为对 VLAN10 的 PC2 的设置截图。

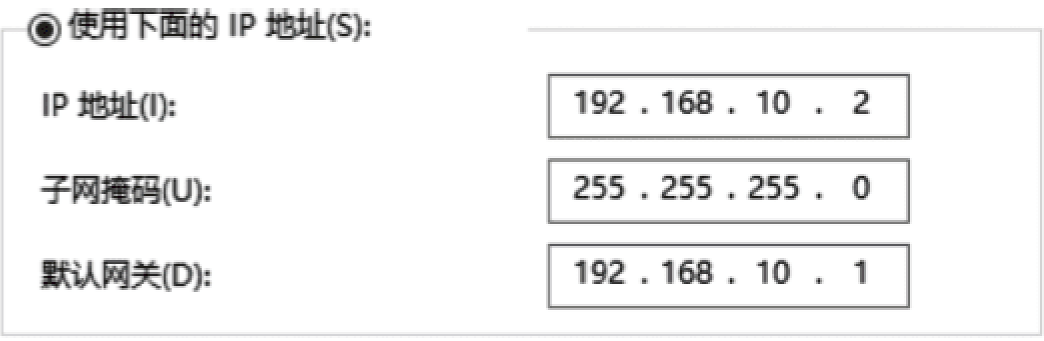


图 6.18 PC2 的设置截图

2. 单臂路由组网配置

当交换机上有多个 VLAN 时，利用多个路由器接口实现 VLAN 间路由选择变得不可能，因为一个路由器没有足够多的以太网接口，并且占用多接口成本将会很高。为此可以使用单臂路由的组网方案，也就是路由器（或三层交换机）的一个快速以太网口和交换机的主干道接口相连，并将路由器的快速以太网物理接口划分成若干个子接口，每个子接口对应一个 VLAN。同时，路由器接口和交换机主干道接口之间将形成主干道，可以运载多个 VLAN 的信息，如图 6.19 所示。要注意的是，如果采用 802.1Q 的封装形式，是不支持 VLAN1 封装的。

依据图 6.13 组网，路由器通过采用子接口方式实现 VLAN 间路由，配置如下。

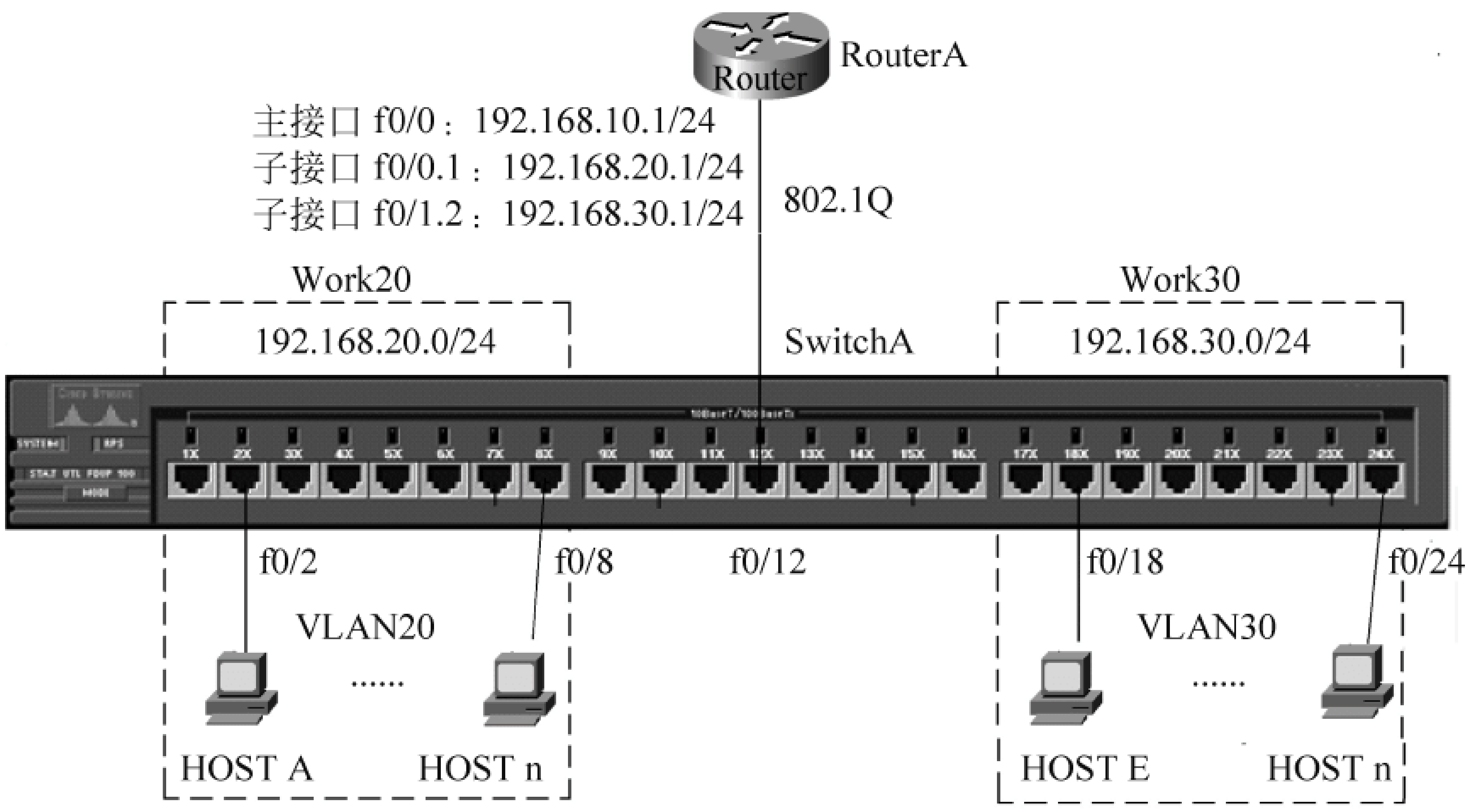


图 6.19 单臂路由

1）路由器 RouterA（可选 cisco2188）配置

```
Router＞enable
Router＃config terminal
Router(config)＃hostname RouterA
RouterA(config)＃interface f0/0                    !进入 f0/0 接口，并配置 IP 地址
RouterA(config－if)＃ip address 192.168.10.1 255.255.255.0
RouterA(config－if)＃no shutdown
RouterA(config－subif)＃interface f0/0.1           !进入子接口 f0/0.1，并配置 VLAN20、协议和 IP
RouterA(config－subif)＃encapsulation dot1q 20     !封装格式为 802.1q，可传输 VLAN20 的帧
RouterA(config－subif)＃ip address 192.168.20.1 255.255.255.0     !子接口 f0/0.1 的 IP 地址
RouterA(config－subif)＃interface f0/0.2           !进入子接口 f0/0.1，并配置 VLAN30、协议和 IP
RouterA(config－subif)＃encapsulation dot1q 30     !封装格式为 802.1q，可传输 VLAN30 的帧
RouterA(config－subif)＃ip address 192.168.30.1 255.255.255.0     !子接口 f0/0.2 的 IP 地址
RouterA(config－subif)＃end
Routera＃show interface f0/0.1              !查看 f0/0.1 子接口是否启动及对应的 VLAN 配置信息
routerA＃show ip router                     !查看路由信息，可以得知 f0/0 设置的子接口数等信息
```

2）交换机 switchA（可选 cisco2960-24TT）配置

```
switch＞enable
switch＃configure terminal
switch(config)＃hostname switchA
switchA(config)＃vlan 20                                  !进入 VLAN 20
switchA(config－vlan)＃name work20                         !创建 VLAN20 的名称为 work20
SwitchA(config－vlan)＃vlan 30                             !进入 VLAN 30
SwitchA(config－vlan)＃name work30                         !创建 VLAN30 的名称为 work30
SwitchA(config－vlan)＃interface range f0/2－8             !进入接口组(f0/2～f0/8)配置
SwitchA(config－if－range)＃switchport mode access         !将该组接口设置为 access 接口模式
SwitchA(config－if－range)＃switchport access vlan 20      !将该组接口加入到 VLAN20
SwitchA(config－if－range)＃interface range f0/18－24      !进入接口组(f0/18～f0/24)配置
SwitchA(config－if－range)＃switchport mode access         !将该组接口设置为 access 接口模式
SwitchA(config－if－range)＃switchport access vlan 30      !将该组接口加入到 VLAN30
SwitchA(config－if－range)＃interface f0/12                !进入接口 f0/12 配置
```

```
SwitchA(config-if)#switchport mode trunk              !将该接口配置为中继接口
switchA(config-if)#switchporth trunk allowed vlan add 10、20        !通过该接口透传的 VLAN
switchA(config-if)#end
```

习题

1. 交换机工作在 OSI 第几层？主要功能是什么？

2. 为什么说交换机通过学习，可以获得源 MAC 地址？源 MAC 地址是否是指同一局域网的其他主机的 MAC 地址？

3. 网桥是如何知道它是应该转发还是应该过滤掉数据包的？

4. 网桥的路由策略有哪几种？

5. 三层交换机能否解决广播域的问题？为什么？

6. 分析三层交换机的原理，并说明它和二层交换机、路由器有哪些区别。

7. 交换机都有哪些基本状态？是如何进入的？

8. 根据图 6.17，完成对路由器和交换机有关 VLAN 间路由的完整配置，包含对各个接口的设置及 VLAN 划分等。也可以通过模拟器或网络机房实验来完成。

9. 参考图 6.19，设计一个简单的网络拓扑结构，并完成单臂路由配置。

第7章 数据链路层协议

HDLC 是一种面向位的控制协议，它对任何一种比特流均可以实现透明传输；PPP 是面向字符的控制协议，是在 SLIP(Serial Line IP)的基础上发展而来的，支持同步和异步串行连接。HDLC 和 PPP 是组建广域网的最常用链路层技术；属于 IEEE 802.3 标准的以太网支持同轴电缆、双绞线和光纤，是组建局域网的最常用链路层技术。本章主要介绍 HDLC、PPP 和以太网技术，以及它们的相关配置。

7.1 高级数据链路控制规程

高级数据链路控制规程(High-Level Data Link Control procedures，HDLC)协议族中的协议都是运行于同步串行线路之上，HDLC 传输控制功能与处理功能分离，具有较大灵活性。

7.1.1 HDLC 技术

目前网络设计普遍使用 HDLC 作为数据链路管制协议，因为它的特点比较明显：不依赖于任何一种字符编码集；数据报文可透明传输，用于实现透明传输的“0”插入法，易于硬件实现；全双工通信，不必等待确认便可连续发送数据，有较高的数据链路传输效率；所有帧均采用 CRC 校验，对信息帧进行编号，可防止漏收或重发，传输可靠性高。

1. HDLC 帧格式

HDLC 是 ISO 开发的，它规定了使用帧字符和校验和的同步串行链路的数据封装方法，规定数据帧中所传输的数据的位数是任意的，不必是字节的整倍数，其帧格式如图 7.1 所示。

	1或2字节	1或2字节	≥0字节	2字节	
01111110	地址	控制	数据	FCS	01111110

图 7.1 HDLC 帧格式

HDLC 的地址字段的内容取决于所采用的操作方式，有主站、从站、组合站之分。命令帧中的字段携带的是对方站的地址，而响应帧中的地址字段所携带的是本站的地址。

CRC 帧校验序列码对从地址开始直至数据字段的内容进行校验。

控制字段用来实现 HDLC 协议的各种控制信息，并标识本帧的类型，根据该字段的内容不同，长度为 1 字节或 2 字节。根据该字段的内容不同，又可分为信息帧、监控帧和无编号帧。

1) 信息帧

信息帧(Ⅰ帧)用于传送有效信息或数据，Ⅰ帧以控制字第一位为“0”来标志。HDLC

允许发送方连续发送多个帧。其中，字段 N(S)代表待发送的帧编号，而字段 N(R)的内容是期待接收的对方下一帧编号，代表对对方已发送过来的帧的确认。

2）监控帧

监控帧(S 帧)用来对通信链路进行控制、管理。其中根据类型字段的不同内容，指示一方怎样解释后面的 N(R)字段。此类型的帧用来实现简单的流控和检错重发。

3）无编号帧

无编号帧(U 帧)因其帧中控制字段不含发送帧编号字段 N(S)和确认帧编号 N(R)而得名。无编号帧主要用来提供各种附加的链路控制命令和响应功能。

最后是 1 字节的帧定界符，标识此帧的结束。

在标准 HDLC 协议格式中我们可以看到，它没有包含标识所承载的上层协议信息的字段，所以在链路层封装标准 HDLC 协议的单一链路上只能承载单一的网络层协议。HDLC 可以使用全双工通信，同时允许发送多帧而只进行单次确认，因此可以达到较高速率。

2. HDLC 连接过程

在实际工作中，两个远程路由器之间是不能直接相连的，因为它们都是 DTE 设备，必须通过数据服务单元(DSU)/信道服务单元(CSU)等这样的局端设备互连，依靠局端设备为两端的路由器提供用于同步的时钟。

1）链路连接阶段

建立数据链路连接阶段。当网络层向链路层发出连接请求时，链路层的发送端向接收端发出置正常响应模式(SNRM)无编号帧，若接收端准备就绪，则发出无编号帧确认(UA)表示同意建立数据链路连接，此时，链路连接就建立好了。

2）传递数据阶段

链路连接建立好后，发送端开始按照某种流量控制策略发送信息帧。如果采用滑动窗口流量控制，那么一次允许连续发送多帧而无需对方应答。接收端收到信息帧后，通过帧校验序列来检验接收的数据是否正确。若正确则发出确认监督帧，否则发出否认监督帧。一般当接收端收到一正确的信息帧后，不急于发出确认监督帧，继续接收后面的信息帧。

3）正常响应方式

正常响应方式(Normal Response Mode，NRM)是一种非平衡数据链路操作方式，有时也称非平衡正常响应方式。该操作方式适用于面向终端的点到点或一点与多点的链路。这种操作方式，传输过程由主站启动，从站只有收到主站某个命令帧后，才能作为响应向主站传输信息。响应信息可以由一个或多个帧组成，若信息由多个帧组成，则应指出哪一个是最后一帧。主站负责管理整个链路，且具有轮询、选择从站及向从站发送命令等权利。

4）异步响应方式

异步响应方式(Asynchronous Response Mode，ARM)也是一种非平衡数据链路操作方式，与 NRM 不同的是，ARM 下的传输过程由从站启动。从站主动发送给主站的一个或一组帧中可包括信息，也可以是仅以控制为目的而发的帧。在这种操作方式下，由从站来控制超时和重发。该方式对采用轮询方式的多站链路来说是必不可少的。

5）异步平衡方式

异步平衡方式(Asynchronous Balanced Mode，ABM)是一种允许任何节点来启动传输的操作方式。为了提高链路传输效率，节点之间在两个方向上都需要较高的信息传输量。

在这种操作方式下任何时候任何站都能启动传输操作，每个站既可作为主站又可作为从站，每个站都是组合站。各站都有相同的一组协议，任何站都可以发送或接收命令，也可以给出应答，并且各站对差错恢复过程都负有相同的责任。

6）帧定界

由于HDLC数据帧中所传输数据的位数是任意的，因此，HDLC是通过约定的位模式进行帧的定界，而不是靠使用特殊定义的字符来界定帧的开始和结束，故称为“面向位”的同步规程。HDLC是使用特殊模式的二进制串“01111110”来标识帧头和帧尾的。如果用户的数据中也出现了这种模式的二进制“01111110”，协议会认为数据帧已结束，从而提前结束数据的接收，导致协议失败。为了避免发生这种现象，发送端在发送数据时，如果用户的数据中出现了二进制串“11111”，则协议会在5个“1”之后自动插入一个“0”；接收端在收到数据的时候，如果连续出现5个“1”之后跟随的是“0”，则自动将它去掉。

7.1.2 HDLC配置

1. 确定DTE、DCE

在实际工作中，HDLC几乎不需要配置就可以工作。但是，在实验室环境下，可能需要做额外的配置，由于条件限制，有时不得不将路由器直接相连(称为背靠背连接)。背靠背连接也用于新购置设备的测试。这时，必须规定哪个路由器是DTE，哪个是DCE，由DCE路由器提供时钟，同时还要设置DCE路由器的时钟频率。

其实，在将两个路由器的串行接口用电缆背靠背连接起来时，就已经决定了哪个路由器会充当DCE端。因为路由器可以自动识别串行接口所接入的电缆类型。可以通过观察电缆类型判断哪个路由器会充当DCE端。方法是看连接电缆(如V.35)一端，如果是孔端接头，则此线缆的另一端就是DCE设备；如果是针端接头，线缆另一端所接的就是DTE设备。

同一台路由器可以同时是DCE和DTE设备，这要视该路由器的对应接口接入了什么类型的电缆而定。如：某台路由器的serial0/0是DCE端设备，而serial0/1却是DTE端设备。

如果路由器处于开机状态，也可以使用“show controllers interface”命令查看路由器某接口的类型，若是DCE设备，将会显示时钟速率。

2. HDLC配置操作

配置接口封装HDLC命令：

```
encapsulation hdlc
```

默认情况下，接口封装的链路层协议为PPP。只有当接口工作在同步方式下时，才能封装HDLC。当接口封装了SLIP时，接口的物理属性不能被修改为同步模式。此时，必须先将接口的链路层封装改为PPP后，才能将接口属性改为同步模式。

【例1】 图7.2给出的是Cisco路由器串行连接，要求写出路由器a、b配置HDLC的过程。

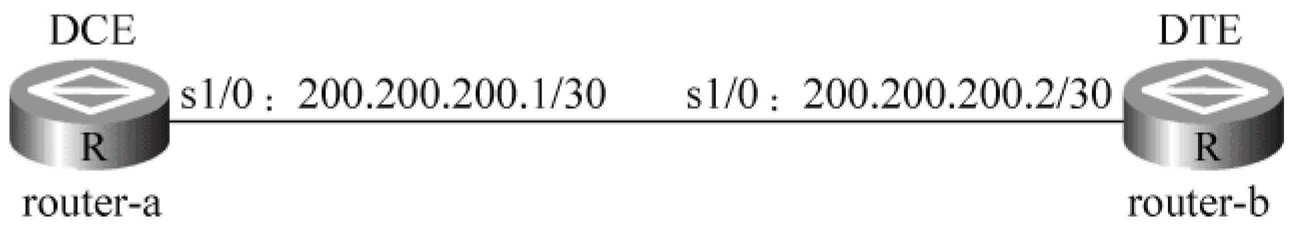


图7.2 配置点到点串行连接

答：(1) router-b 作为 DTE 端，配置过程如下。

```
router-b#configure terminal
router-b(config)#interface serial 1/0                          !进入 s1/0
router-b(config-if)#encapsulation hdlc                         !封装 HDLC
router-b(config-if)#ip address 200.200.200.2 255.255.255.252   !配置 IP 地址
router-b(config-if)#no shutdown                                !激活该接口
router-b(config-if)#end
```

(2) 作为 DCE 端的 router-a，需要使用 clock rate 命令设置 DCE 端的时钟，范围为 1200～8 000 000，单位是 bps。router-a 配置过程如下。

```
router-a#configure terminal
router-a(config)#interface serial 1/0
router-a(config-if)#encapsulation hdlc                         !封装 HDLC
router-a(config-if)#ip address 200.200.200.1 255.255.255.252
router-a(config-if)#clock rate 2000000                         !设置 DCE 端的时钟
router-a(config-if)#no shutdown
router-a(config-if)#end
```

配置完成后，可以使用命令 show interface serial 1/0 检查接口状态，显示了物理接口及协议状态。如果物理接口状态为 down，表明物理连接有问题，没有收到任何信号；如果物理接口状态为 up，协议状态字段为 down，则表明物理连接正常，但是数据链路层工作有问题。一般在实际环境中不需要配置时钟频率，因为通信服务商会利用其 DCU/CSU 提供时钟。

7.2 点到点协议

点到点协议(Point-to-Point Protocol，PPP)由链路控制协议族(Link Control Protocol，LCP)、网络层控制协议族(Network Control Protocol，NCP)以及 PPP 扩展协议族组成，PPP 支持同步和异步串行连接，支持多种网络层协议。

7.2.1 PPP 技术

PPP 协议是广域网上应用最为广泛的协议之一，它的优点在于简单、具备用户验证能力，可以解决 IP 分配等。家庭拨号上网就是通过 PPP 在用户端和运营商的接入服务器之间建立通信链路。目前，典型的应用是在 ADSL 接入方式中，PPP 与其他的协议共同派生出了符合宽带要求的新协议，如 PPPoE(PPP over Ethernet)，PPPoA(PPP over ATM)等。

1. PPP 协议的特点和帧格式

PPP 协议的特点为：PPP 协议是数据链路层协议；支持点到点的连接；物理层可以是同步电路或异步电路；具有各种 NCP 协议，如 IPCP(IP 控制协议)，IPXCP(网际数据包交换控制协议)，更好地支持了网络层协议；具有验证协议 PAP/CHAP，更好地保证了网络的安全性。

PPP 帧格式以 HDLC 帧格式为基础，做了很少的改动。两者的主要区别是：PPP 是面向字符的，而 HDLC 是面向位的。PPP 在点到点串行线路上使用字符填充技术，所有帧的

大小都是字符的整数倍。图 7.3 中给出了 PPP 的帧格式，各字段说明如下。

			1~2字节	<150字节	2或4字节	
01111110	11111111	00000011	协议	数据	FCS	01111110

图 7.3　PPP 帧的格式

第一字段和最后字段相同，PPP 帧是以标准 HDLC 标志字节(01111110)开始和结束的。

第二字段是地址字段，默认情况下，被固定设成二进制数 11111111，因为点到点线路的一个方向上只有一个接收方。

第三字段是控制字段，默认情况下，被固定设成二进制数 00000011。

因为默认情况下，地址字段、控制字段总是常数，因此，这两部分实际可以省略不要，当然需要通过链路控制协议(Link Control Protocol，LCP)进行协商。

第四字段为协议字段，用来标明后面携带的是什么类型的数据，其默认大小为 2 字节，但如果是 LCP 包，则可以是 1 字节。

第五字段是数据字段，其长度可变，默认最大长度为 1500 字节。

第六字段是校验和字段，通常情况下是 2 字节，但也可以是 4 字节。

2. PPP 结构和工作过程

1) PPP 分层结构

PPP 协议作为一种提供在点到点链路上封装、传输网络层数据包的数据链路层协议，处于 OSI 参考模型的第二层，主要用来在支持全双工的同、异步链路上进行点到点之间的数据传输，也是许多路由器串口默认数据链路层封装协议。

如图 7.4 所示，PPP 处在数据链路层，支持各种类型的硬件，包括 EIA/TIA232、EIA/TIA449、EIA/TIA530、V.35、V.21 等。只要是点到点类型的线路都可以运行 PPP。在数据链路层，由 NCP 为不同的协议提供服务。这里的 NCP 相当于以太网数据链路层的 LLC 子层。

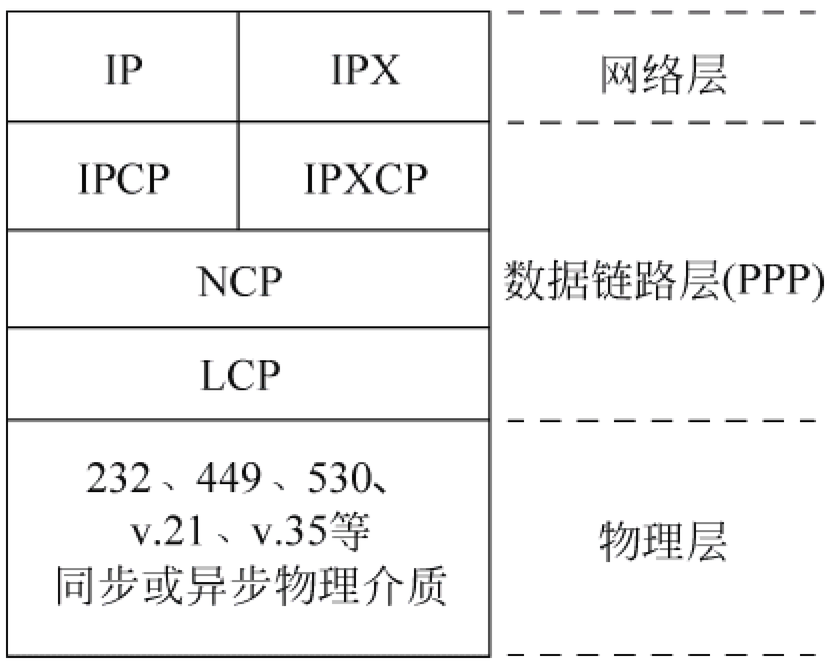


图 7.4　低层协议结构中的 PPP

PPP 主要由两类协议组成：链路控制协议族和网络层控制协议族。链路控制协议主要用于建立、拆除和监控数据链路，NCP 主要用于协商在该数据链路上所传输的数据包的格式与类型。同时，PPP 还提供了用于网络安全方面的验证协议族(PAP 和 CHAP)。

(1) PPP 扩展协议族：提供对 PPP 功能的进一步支持。在串行链路上封装 IP 数据报的方法。PPP 既支持数据为 8 位和无奇偶检验的异步模式(如计算机上的串行接口)，也支持面向位的同步连接。

(2) LCP 是建立、配置及测试数据链路的链路控制协议。它允许通信双方进行协商，以确定不同的选项。

(3) NCP 是针对不同网络层协议的网络控制协议。NCP 为 PPP 提供上层服务接口。当 LCP 将链路建立好以后，PPP 开始根据不同用户的需要，配置上层协议所需的环境。NCP 协议使 PPP 可以支持 IP、IPX 等多种网络层协议及 IP 地址的自动分配。

2）PPP 工作过程

用户在入网操作时，可利用电话线通过调制解调器接入 Internet 服务提供商（Internet Service Provider，ISP）路由器。ISP 局端设备收到用户的接入呼叫后，就分配给该用户一个临时的 IP 地址，使该用户的 PC 通过 ISP 路由器成为接在 Internet 上的主机，并使用 Internet 提供的各种服务。此时，用户计算机中使用 TCP/IP 的“客户进程”与 ISP 路由器中的“选路进程”建立起一个 TCP/IP 连接，用户正是通过这个连接与 Internet 通信的。下面是对过程的描述。

（1）当一个 PC 终端拨号用户发起一次拨号后，此 PC 终端首先通过调制解调器呼叫远程访问服务器，如提供拨号服务的路由器。

（2）当路由器上的远程访问模块应答了这个呼叫后，就建立起一个初始的物理连接。

（3）接下来，PC 终端和远程访问服务器之间开始传送一系列经过 PPP 封装的 LCP 分组，用于协商选择将要采用的 PPP 参数。

（4）如果上一步中有一方要求认证，接下来就开始认证过程。如果认证失败，则链路被终止，关闭物理链路回到空闲状态。如果认证成功则进行下一步。

（5）通信双方开始交换一系列的 NCP 分组来配置网络层。对于上层使用过程是由 IP/TCP 来完成的。

（6）当 NCP 配置完成后，双方的逻辑通信链路就建立好了，双方可以开始在此链路上传送上层数据。

（7）当数据传送完成后，一方会发起断开连接的请求。这时，首先使用 NCP 来释放网络层的链接，归还 IP 地址；然后利用 LCP 来关闭数据链路层连接；最后，使双方的通信设备或模块关闭，物理链路回到空闲状态。

3. PPP 协商

1）PPP 协商流程

PPP 协商流程如图 7.5 所示，它分为 5 个阶段：Dead、Establish、Authenticate、Network 和 Terminate 阶段，在不同的阶段进行不同协议的协商。只有在前面的协议协商出结果后，才能转入下一个阶段，进行下一个协议的协商。

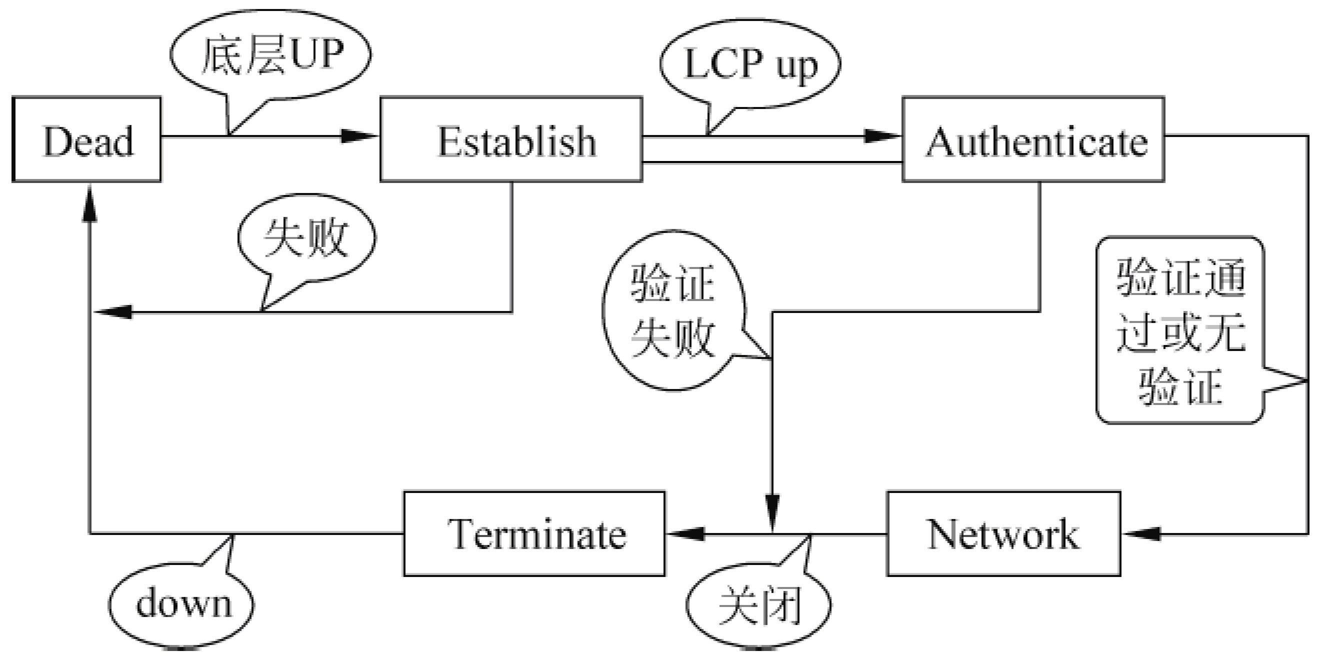


图 7.5　PPP 协商流程

（1）当物理层不可用时，PPP 链路处于 Dead（无效）阶段，链路必须从这个阶段开始和结束。当物理层可用时，PPP 在建立链路之前首先进行 LCP 协商，协商内容包括工作方式、验证方式和最大传输单元等。

(2) LCP 协商过后就进入 Establish(建立)阶段,此时 LCP 状态为 Opened(开启),表示链路已经建立。

(3) 如果配置了验证(远端验证本地或者本地验证远端)就进入 Authenticate(验证)阶段,开始 CHAP 或 PAP 验证。

(4) 如果验证失败,进入 Terminate(终止)阶段,拆除链路,LCP 状态转为 Down;如果验证成功就进入 Network(网络)协商阶段(NCP),此时 LCP 状态仍为 Opened,而 IPCP 状态从 Initial(初始)转到 Request(请求)。

(5) NCP 协商支持 IPCP 协商,IPCP 协商主要包括双方的 IP 地址。通过 NCP 协商来选择和配置一个网络层协议。当选中的网络层协议配置成功后,该网络层协议就可以通过这条链路发送报文。

(6) PPP 链路将一直保持通信,直至有明确的 LCP 或 NCP 帧关闭这条链路,或发生了某些外部事件(例如,用户的干预)。

2) LCP 的协商选项与链路协商过程

(1) LCP 的协商选项。LCP 规定了链路建立、维护以及拆除。LCP 用来在通信链路建立初期,在通信双方之间协商功能选项。表 7.1 列出其中主要的选项,包括身份认证、链路压缩、回叫、多链路捆绑等。

表 7.1 LCP 的协商选项

特征	解释	协议
身份认证	链路建立成功前要求提供正确的密码	PAP,CHAP
链路压缩	在宽带有限的链路上,提供对数据的压缩功能	MPPC(点对点压缩)等
回叫	由被叫方重新呼叫原呼叫方发起	Cisco Callback,MS Callback
多链路捆绑	根据需要进行多链路捆绑、负载均衡	MP

PPP 协议运行在点到点串行链路上,为了提高数据发送效率,可以采用对数据进行压缩后再传送的方法,称为链路压缩。

回叫又称为回拨,是指通信一方拨号到另一方后,由另一方断开拨号连接并进行反向的拨号。回叫有更安全的优点。因为乙方在回叫之前可以验证对方是否为合法用户,可以用口令数据库的方法。

LCP 的多链路捆绑(MultiLink PPP,MP)选项通过将通信两端之间的多条通信链路捆绑成一条虚拟的链路而达到扩充链路可用带宽的目的。LCP 的多链路捆绑可以在多种类型的物理接口上实现,包括异步串行接口、同步串行接口、BRI、PRI 等。MP 允许将报文分片,分片将从多个点对点链路上送到同一个目的地。

(2) MP 方式下链路协商过程。首先和对端进行 LCP 协商,协商过程中,除了协商一般的 LCP 参数外,还验证对端接口是否也工作在 MP 方式下。如果对端不工作在 MP 方式下,则在 LCP 协商成功后,进行一般的 NCP 协商步骤,不进行 MP 捆绑。

然后对 PPP 进行验证,得到对方的用户名。如果在 LCP 协商中得知对端也工作在 MP 方式下,则根据用户名找到为该用户指定的虚拟接口模板,并以该虚拟模板的各项 NCP 参数(如 IP 地址等)为参数进行 NCP 协商,物理接口配置的 NCP 参数不起作用。NCP 协商通过后,即可建立 MP 链路,用更大的带宽传输数据。

一个 PPP 通道如果在 LCP 中协商了如下参数，则它能被绑定为 MP 的一个子通道。

MRRU(Maximum Received Reconstructed Unit)：最大接收重组单元，与普通 PPP 中的 MRU 参数类似。

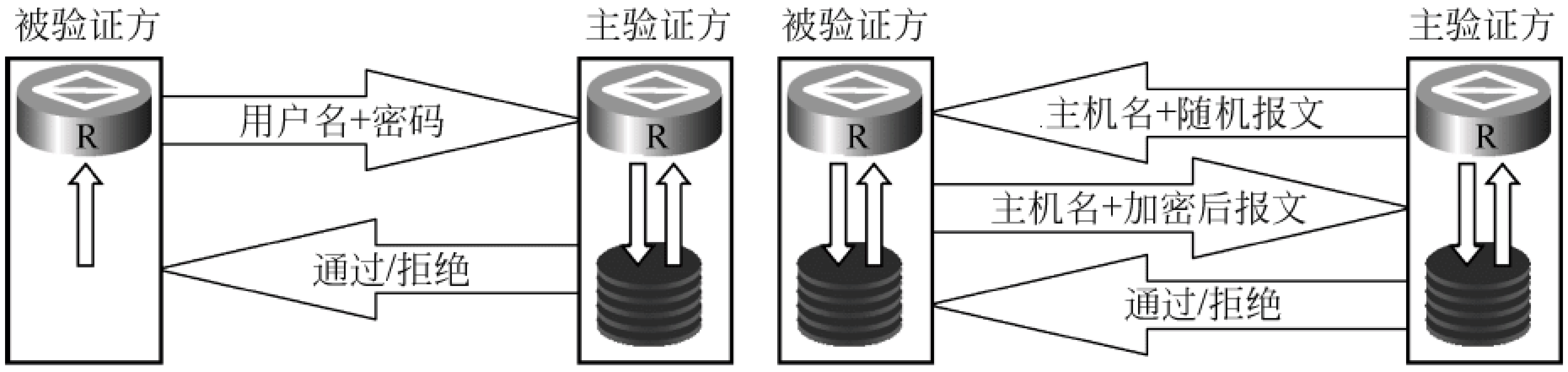
SSNHF(Short Sequence Number Header Format)：短序列号 MP 报文头，这是可选参数。

ED(Endpoint Discriminator)：终端描述符。是唯一标志一个网络实体(路由器、主机等)的字符串。只有终端描述符相同的 PPP 通道可以绑定到同一个 MP。

如 PPP 配置了用户验证功能，则 MP 子通道在验证通过后，就要把自己绑定到一个 MP 上。用于绑定的标志有两个：用户名和终端描述符。

4. PAP/CHAP 验证

密码认证协议(Password Authentication Protocol，PAP)是简单认证方式，采用明文传输，验证只在开始联机时进行。质询握手身份验证协议(Challenge-Handshake Authentication Protocol，CHAP)是要求握手验证方式，安全性较高，采用密文传送用户名，主验方和被验方两边都有数据库，要求双方的用户名互为对方的主机名，即本端的用户名等于对端的主机名，且口令相同。PAP/CHAP 验证示意如图 7.6 所示。



• PAP是两次握手验证协议，口令以明文传送，被验证方首先发起验证请求。

(a) PAP验证

• CHAP是三次握手验证协议，不发送口令，主验证方首先发起验证请求，安全性比PAP高。

(b) CHAP验证

图 7.6　PAP/CHAP 验证示意

1) PAP 验证

PAP 验证为两次握手验证，口令为明文，如图 7.6(a)所示，PAP 验证的过程如下。

(1) 被验方先发起连接，将 username 和 password 一起发给主验方。主验方收到被验方 username 和 password 后，在数据库中进行匹配，并回送 ACK(Acknowledge)或 NAK(Not Acknowledge)。

(2) 如正确，主验方则会给对端发送 ACK 报文，通告对端已被允许进入下一阶段协商。否则，发送 NAK 报文，通告对端验证失败。此时，并不会直接关闭链路。只有当验证不通过次数达到一定值(默认为 4)时，才会关闭链路，来防止因误传、网络干扰等造成不必要的 LCP 重新协商过程。

PAP 的特点是在网络上以明文的方式传递用户名及口令，如在传输过程中被截获，便有可能对网络安全造成极大的威胁。因此，它适用于对网络安全要求相对较低的环境。

2) CHAP 验证

CHAP 验证为三次握手验证，口令为密文(密钥)，如图 7.6(b)所示，验证过程如下。

(1) 验证方向被验证方发送一些随机产生的报文，并同时将本端的主机名附带上一起发送给被验证方。

(2) 主被验证方接到对端对本端的验证请求(Challenge)时,便根据此报文中验证方的主机名和本端的用户表查找用户口令字,如找到用户表中与验证方主机名相同的用户,便利用接收到的随机报文、此用户的密钥用 MD5 算法生成应答(Response),随后将应答和自己的主机名送回。

(3) 验证方接到此应答后,利用对端的用户名在本端的用户表中查找本方保留的口令字,用本方保留的口令字(密钥)和随机报文用 MD5 算法得出结果,与被验证方应答比较,根据比较结果返回相应的结果(ACK 或 NAK)。

7.2.2 PPP 配置

1. PPP 基本配置

1) Quidway 命令

(1) 封装 PPP: **link-protocol ppp**

它指定一个广域网口的封装类型为 PPP。默认情况下,封装的链路层协议即为 PPP。

(2) 设置验证类型: **ppp authentication-mode**{pap|chap}

它指定验证方式,可选的验证方式为 PAP 和 CHAP。需要注意的是:验证是单向的,配置这条命令的一方作为验证方来验证对方。如果通信的双方都要验证对方,则双方都应配置此命令。

(3) 对端所需的用户名和密码: **local-user** username **password** {simple|cipher} password

它配置验证所需的用户名和口令。simple 表示以明文的方式显示后面的口令,cipher 表示以加密的方式显示后面的口令,最后面的 password 就是设置的口令。

2) sisco 命令

(1) 封装 PPP: **encapsulation ppp**

(2) 设置验证类型: **ppp authentication pap**

或 ppp authentication pap chap

或 ppp authentication chap pap

默认情况下,Cisco 路由器将接受 CHAP 作为认证协议。在客户端希望执行 PAP,但接入服务器可执行 PAP 或 CHAP。

(3) 对端所需的用户名和密码: **ppp pap sent-username** username **password** password

说明:用户名和密码与接收路由器上的用户名和密码匹配;路由器两端必须都进行 PPP 封装;创建用户名时注意与对方设备名一致,但设置的用户密码可以不同;路由器两端都要进行 PAP 验证的配置,用于验证发送方的用户名及密码。

例如:在 routerA 的 serial 0/0 配置 pap,对端用户名:routerB,密码:abc。配置如下:

```
routerA(config)#username routerB password abc                 !配置对端用户名和密码
routerA(config)#interface serial 0/0
routerA(config-if)#encapsulation ppp                           !在 s0/0 封装 PPP 协议
routerA(config-if)#ppp authentication pap                      !配置 PAP 认证
routerA(config)#ppp pap sent-username routerA passwoed abc     !配置对端所需的用户名和密码
```

【例 2】 如图 7.7 所示,路由器 routerA 和 routerB 之间采用 PPP 协议,routerA 的 Serial2/0 与对方 routerB 的 Serial3/0 实现互连,要求路由器 routerA 用 PAP 方式验证路

由器 routerB。

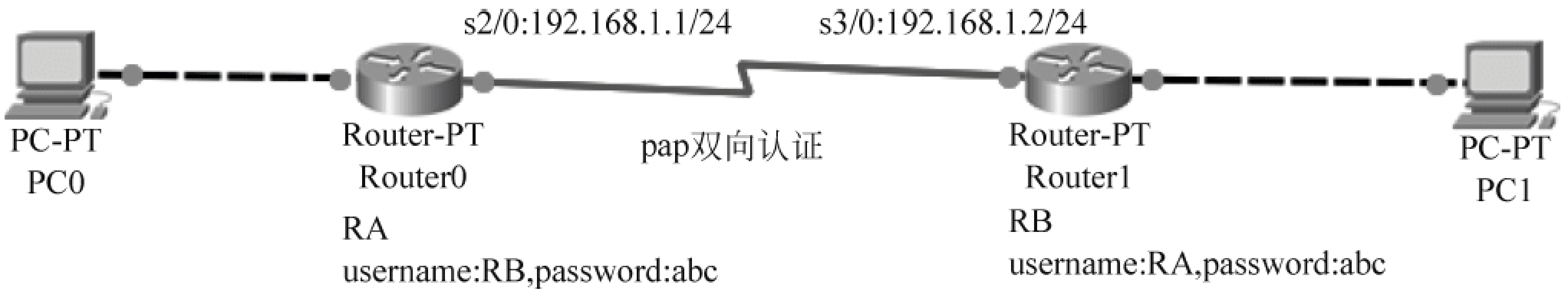


图 7.7　PPP 配置实例

答：(1) 配置路由器 RA。

```
Router>enable
Router#configure terminal
Router(config)#hostname RA
RA(config)#username RB password abc                    !配置对端用户名和密码
RA(config)#interface s2/0
RA(config-if)#ip address 192.168.1.1 255.255.255.0
RA(config-if)#encapsulation ppp                        !封装 PPP
RA(config-if)#ppp authentication pap                   !设置 PAP 验证方式
RA(config-if)#ppp pap sent-username RA password abc    !送给对方的用户名及密码
RA(config-if)#clock rate 64000                         !配置串口 DCE 同步时钟
RA(config-if)#no shutdown
RA(config-if)#exit
```

(2) 配置路由器 RB。

```
Router>enable
Router#configure terminal
Router(config)#hostname RB
RB(config)#username RA password abc                    !配置对方用户名和密码
RB(config)#int s3/0
RB(config-if)#ip address 192.168.1.2 255.255.255.0
RB(config-if)#encapsulation ppp
RB(config-if)#ppp authentication pap
RB(config-if)#ppp pap sent-username RB password abc    !送给 RA 的用户名及密码
RB(config-if)#no shutdown
RB(config-if)#exit
```

(3) 查看路由器 RB 配置。

```
RB#show interface s3/0
```

在路由器上执行以上命令后，就会显示如图 7.8 所示的文字信息，即路由器 RB 串口 s3/0 部分截图内容如下：

Serial3/0 激活，线路协议已接通(线路两端协议是一致的)；

硬件地址是 hd64570；

互联网地址是 192.168.1.2/24；

MTU 为 1500 字节，BW 为 128kbit，延时为 20000μs；

接口可靠性(reliability)为 255/255，其中，发包接口的负载(txload)为 1/255，收报接口

```
RB>show interface s3/0
Serial3/0 is up, line protocol is up (connected)
  Hardware is HD64570
  Internet address is 192.168.1.2/24
  MTU 1500 bytes, BW 128 Kbit, DLY 20000 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation PPP, loopback not set, keepalive set (10 sec)
  LCP Open
  Open: IPCP, CDPCP
```

图 7.8 查看路由器 RA 串口部分截图

的负载(rxload)为 1/255;

接口封装协议 PPP,接口回环没有设置,存活定时器(10s);

LCP 已开启;

开放:IPCP,CDPCP。

2. MP 协议配置

在配置 MP 之前,需要先完成虚拟接口模板的配置。被绑定在虚拟接口模板下的接口,首先还必须配置和对端进行双向验证(CHAP 或 PAP)。Quidway 主要命令如下。

(1) 配置封装 PPP 的接口工作在 MP 方式

命令:**ppp mp**

在接口配置模式,执行该命令可以完成封装 PPP 的接口工作在 MP 方式。

(2) 建立虚拟接口模板与 MP 用户的联系

命令:**ppp mp user** user-name **bind virtual-template** number

在全局配置模式下,执行该命令可以建立虚拟接口模板与 MP 用户的对应关系。

完成以上配置后,MP 基本配置已经完成。用户可以根据自己的实际需要进行其他针对 MP 的可选参数配置,如配置 MP 最大绑定链路数。其中 user-name 为 MP 用户的用户名,number 为绑定的虚拟模板数号。

(3) 配置 MP 最大绑定链路数

命令:**ppp mp max-bind** binds

在全局配置模式下,采用该命令可以完成 MP 最大绑定链路数的配置。binds 取值范围为 1~100,在实际配置本参数时,应考虑需求,因为捆绑的链路过多,会降低系统性能。

【例 3】 如图 7.9 所示,3 个路由器 a、b 和 c 通过 DDN 互连,假设 Router-a 的 E1 口取出 4 个 B 信道(64Kbps)的接口名分别为 serial0、serial1、serial2、serial3;Router-b 的 E1 口取出两个 B 信道的接口名分别为 serial0、serial1;Router-c 的 E1 口取出两个 B 信道的接口名分别为 serial0、serial1。要求通过 MP 绑定链路,将 Router-a 的两个 B 信道绑定到 Router-b 上,另外两个 B 信道绑定到 Router-c 上。

答:1) 配置路由器 Router-a

(1) 增加两个用户 Router-b 和 Router-c。

```
[Quidway]local-user router-b password simple b123        !配置验证的用户名和明文口令
[Quidway]local-user router-c password simple c123
```

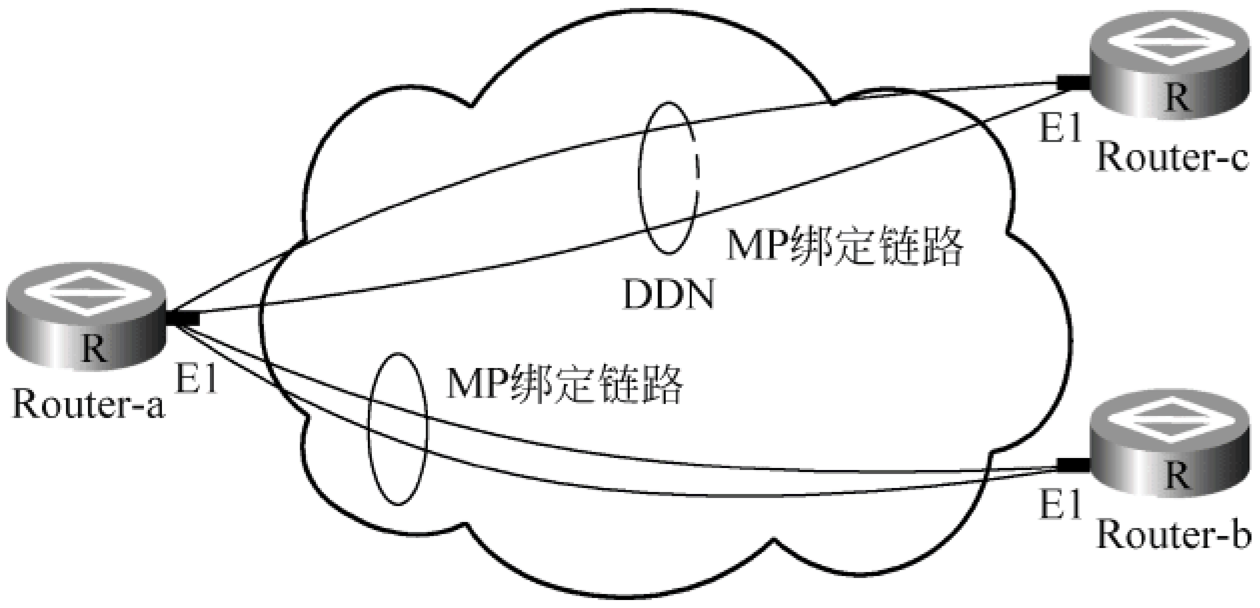


图 7.9 MP 典型配置举例

(2) 为增加用户指定虚拟接口模板,将使用该模板的 NCP 信息进行 PPP 协商。

```
[Quidway]ppp mp user rourter-b bind virtual-template 1   !接口模板 1 与 Rourter-b 建立联系
[Quidway]ppp mp user rourter-c bind virtual-template 2   !接口模板 2 与 Rourter-c 建立联系
```

(3) 配置虚拟接口模板 IP 地址。

```
[Quidway]interface virtual-template 1                    !进入虚拟接口模板 1
[Quidway-Virtual-Template1]ip address 200.20.2.1 255.255.255.0
[Quidway]interface virtual-template 2                    !进入虚拟接口模板 2
[Quidway-Virtual-Template2]ip address 200.20.3.1 255.255.255.0
```

(4) 将接口 serial0、serial1、serial2、serial3 加入 MP 通道,以下只给出 serial0 的配置,其他接口也要作类似的配置。

```
[Quidway]interface serial 0
[Quidway-Serial0]link-protocol ppp                       !接口 serial 0 封装为 PPP
[Quidway-Serial0]ppp mp                                  !完成封装 PPP 的接口工作在 MP 方式
[Quidway-Serial0]ppp authentication-mode pap             !配置 PPP 身份验证模式为 PAP
[Quidway-Serial0]ppp pap local-user router-a password simple a123
                !配置本地被对端以 PAP 方式验证时本地发送的 PAP 用户名 router-a 和口令 a123
```

2) 配置路由器 Router-b

(1) 增加一个用户 Router-a。

```
[Quidway]local-user router-a password simple a123        !配置验证的用户名和明文口令
```

(2) 为增加用户指定虚拟接口模板,将使用该模板的 NCP 信息进行 PPP 协商。

```
[Quidway]ppp mp user router-a bind virtual-template 1    !接口模板 1 与 Rourter-a 建立联系
```

(3) 配置虚拟接口模板的工作参数。

```
[Quidway]interface virtual-template 1                    !进入虚拟接口模板 1
[Quidway-Virtual-Template1]ip address 200.20.2.2 255.255.255.0
```

(4) 将接口 serial0、serial1 加入 MP 通道,以下只给出 serial0 的配置,serial1 也要作同样的配置。

```
[Quidway]interface serial 0
[Quidway-Serial0]link-protocol ppp                       !接口 serial 0 封装为 PPP
```

```
[Quidway－Serial0]ppp mp                                  !完成封装 PPP 的接口工作在 MP 方式
[Quidway－Serial0]ppp authentication－mode pap            !完成封装 PPP 的接口工作在 MP 方式
[Quidway－Serial0]ppp pap local－user router－b password simple b123
                            !配置本地被对端以 PAP 方式验证时本地发送的 PAP 用户名 router－b 和口令 b123
```

3）配置路由器 Router-c

（1）增加一个用户 Router-a。

```
[Quidway]local－user router－a password simple a123       !配置验证的用户名和明文口令
```

（2）为这个用户指定虚拟接口模板，将使用该模板的 NCP 信息进行 PPP 协商。

```
[Quidway]ppp mp user router－a bind virtual－template 1   !接口模板 1 与 Rourter－a 建立联系
```

（3）配置虚拟接口模板的工作参数。

```
[Quidway]interface virtual－template 1                     !进入虚拟接口模板 1
[Quidway－Virtual－Template1]ip address 200.20.3.2 255.255.255.0
```

（4）将接口 serial0、serial1 加入 MP 通道，以下只给出 serial0 的配置，serial1 也要作类似的配置。

```
[Quidway]interface serial 0
[Quidway－Serial0]ppp mp                                  !完成封装 PPP 的接口工作在 MP 方式
[Quidway－Serial0]ppp authentication－mode pap            !完成封装 PPP 的接口工作在 MP 方式
[Quidway－Serial0]ppp pap local－user router－c password simple c123
                            !配置本地被对端以 PAP 方式验证时本地发送的 PAP 用户名 router－c 和口令 c123
```

7.3 以太网

从传统以太网(Ethernet)向交换式以太网过渡是一个质的转变，而从十兆到百兆、千兆以太网的过渡则成为一个量的转变。带宽的不断提高，使得以太网技术永葆活力。

7.3.1 以太网技术

1. 以太网分类

1）传统以太网

我们知道，传统以太网采用的是带冲突检测的载波侦听多路接入(Carrier Sense，Multiple Access with Collision Detection，CSMA/CD)技术，网络中所有主机的收发都依赖于同一套介质，是在共享介质条件下的多点通信，其基本通信规则如下。

（1）若介质空闲，传输；否则，转(2)。

（2）若介质忙，一直监听到信道空闲，然后立即传输。

（3）若在传输中测得冲突，则发出一个短小的人为干扰信号，使得所有站点都知道发生了冲突并停止传输。

（4）发完人为干扰信号，等待一段随机的时间后，再次试图传输，回到(1)重新开始。

传统以太网试图通过网桥来分隔主机，形成多个冲突域，这样单播报文就会被限制在自己的冲突域内，从而减少报文碰撞的发生，但整个局域网仍然是一个广播域。

2）交换式以太网

交换式以太网就是能为用户提供独享的、更高的带宽，它的特点是：平时网络中的所有主机都不连接，当主机需要通信时，通过交换设备连接对端主机，通信完成后断开；使用交换式集线器或交换机等设备组网，物理上或逻辑上均为星形结构；分割了网络的碰撞域，使得网络减少冲突；由于交换设备的智能化，能为用户提供更多的功能，如 VLAN、RMON、流控和网管等；更大程度上满足了网络对于高速、可靠、扩充性好等需求，将局域网由单纯的数据传输网络提升到适合话音、数据、图像等融合的新境界。以下介绍交换式以太网技术。

（1）VLAN 技术，就是将一个交换网络逻辑地划分成若干子网，每一个子网就是一个广播域。逻辑上划分的子网在功能上与传统物理上划分的子网相同，引入 VLAN 技术可以高效地组播控制，便于网络监督和管理，减少对路由器的依赖。IEEE 802.1Q 定义了标准的 VLAN 格式，并保证 VLAN 信息在不同厂家设备间的互通。

（2）信息流优先级技术，使得交换机将优先转发优先级别较高的数据或信息流，从而保证多媒体数据在局域网中以最小的延迟和足够的带宽进行传输。对于交换式以太网，IEEE 802.1p 定义了通过在以太网帧首部添加优先级信息的规则；对于三层交换机和路由器，IPv4 的 TOS 域定义了标识 IP 信息流优先级的标准。

（3）组播技术，是为了解决点到多点通信而发展起来的，发送者只发送一次报文，路由器和交换机自动将报文复制给每一个真正需要接收报文的终端。组播的实现包括了三层组播路由和组播管理协议两个方面。以太网交换机通过侦听 IGMP 报文或二层组播登记协议 GMRP 来识别组播成员，从而保证组播报文正确、有效地传送。

（4）远程监测（RMON），提供给网络管理人员一种比传统的 SNMP 更加有效的网络监测和分析手段，智能的 RMON 代理自动采集并保存历史统计信息，并且不会干扰网络的工作，同时减少了网管工作站的负荷。

（5）生成树协议，允许交换机之间存在冗余链路，交换机通过生成树算法完成对冗余链路的管理，使得只有一条链路工作，其余链路被阻塞。当使用中的链路出现故障时，自动启动的生成树算法将原来阻塞的冗余链路改变为工作状态，保证交换机之间存在正常的通信链路。

2. 以太网帧格式

图 7.10 给出两种以太网的帧格式，最常用的以太网 IP 数据报格式是 RFC 894 封装格式，IEEE 802.3z 是吉比特以太网的技术标准。在帧格式中，各字段内容如下。

（1）目的地址字段：确定帧的接收者，几乎所有的 802.3 网络都采用 6 字节寻址。

（2）源地址字段：标识发送帧的工作站，它和目前地址字段类似，前 3 字节表示由 IEEE 分配给厂商的地址，将烧录在每一块网络接口卡的 ROM 中。通常由制造商为每一网络接口卡分配后 3 字节。

（3）长度字段：用于 IEEE 802.3 的两字节长度字段定义了数据字段包含的字节数。

（4）类型字段：两字节的类型字段仅用于 Ethernet Ⅱ帧。该字段用于标识数据字段中包含的高层协议，例如：类型字段取值为十六进制 0800 的帧将被识别为 IP 协议帧，而类型字段取值为十六进制 8137 的帧将被识别为 IPX 和 SPX 传输协议帧。

在 IEEE 802.3 标准中类型字段被替换为长度字段，因而 Ethernet Ⅱ帧和 IEEE 802.3

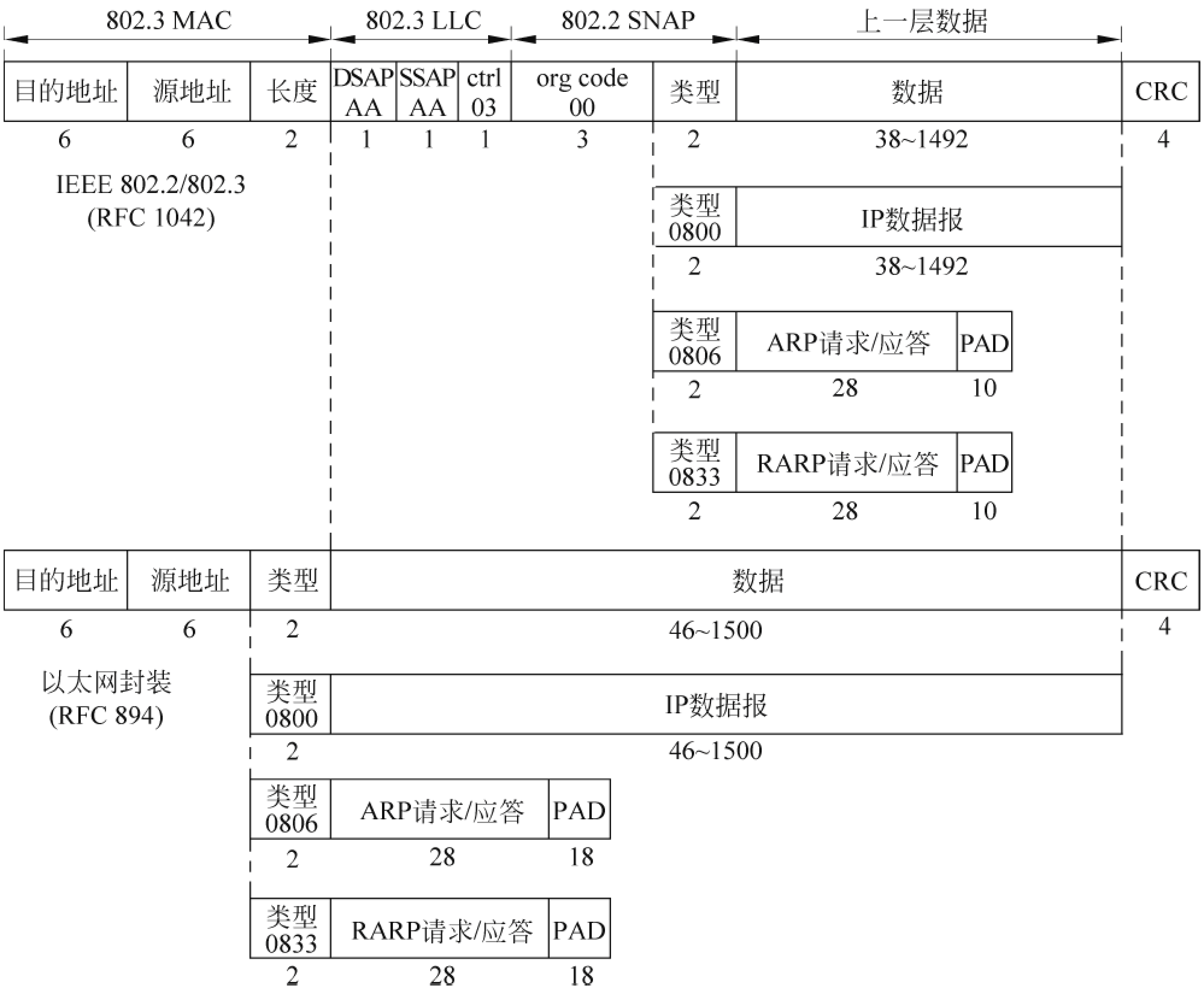


图 7.10　IEEE 802.2/802.3(RFC 1042)和以太网封装格式(RFC 894)

帧之间不能兼容。

DSAP(AA)、SSAP(AA)、ctrl(03)和 org code(00)这些参数一般都是固定的。

(5) 数据字段：数据字段的最小长度必须为 46 字节，以保证帧长至少为 64 字节，这意味着传输 1 字节信息也必须使用 46 字节的数据字段。如果填入该字段的信息少于 46 字节，该字段的其余部分也必须进行填充。数据字段的最大长度为 1500 字节。

(6) CRC 字段：校验和字段。既可用于 Ethernet Ⅱ，又可用于 IEEE 802.3 标准的帧校验序列字段提供一种错误检测机制，每一个发送器均计算一个包括了地址字段、类型/长度字段和数据字段的循环冗余校验(CRC)码。发送器于是将计算出的 CRC 填入 4 字节的 CRC 字段，也称为帧校验或帧检验序列(Frame Check Sequence，FCS)。

3. 以太网连接速率

标准以太网只能支持 10Mbps 的数据传输速率，快速以太网能支持 100Mbps 的数据传输速率，而吉比特以太网则能支持 1000Mbps 的数据传输速率。

1) 标准以太网

传统以太网选择的介质类型绝大多数使用一种共享介质结构。为用户增加带宽可以有两种方法：一种方法是增加网络的总体带宽；另一种方法是减少在同一共享介质线缆段上的设备数量，即采用交换式的以太网设备替代原来的集线器。组建交换式以太网，它可以为每个用户提供独享的 10Mbps，同时还可以配置成全双工工作方式，进一步提升网络的性能。

2）快速以太网

目前一般新建方案都是 100Mbps 以上。也可以将原来标准以太网的速率从 10Mbps 增加到 100Mbps，只需将原有的 10Mbps 集线器或者以太网交换机升级成为快速以太网交换机，用户更换一块 100Mbps 的网卡即可。许多实际运行的网络均存在众多的客户机试图访问同一台服务器的情况，从而在服务器和以太网之间产生瓶颈，为了增强服务器的访问性能，可以通过快速以太网连接以保证访问速度。

3）吉比特以太网

许多汇聚层的以太网交换机均提供千兆接口，用于连接其他的交换机，组成更大的网络，许多支持堆叠功能的以太网交换机也是采用千兆接口实现堆叠功能的。所谓堆叠，是指通过软、硬件的支持，将一组交换机连接起来作为一个对象加以控制的方式，但由于是一种非标准技术，通常不支持各个厂家交换机的混合堆叠。

7.3.2 以太网接口

1. 自协商技术

以太网技术发展到 100Mbps 速率以后，出现了一个如何与原 10Mbps 以太网设备兼容的问题，自协商技术就是为了解决这个问题而制订的。从图 7.11 中可以看出以太网自协商技术的具体应用情况。自协商功能允许一个网络设备将自己所支持的工作模式信息传达给在网络上通信的对端，并接收对端可能传递过来的相应信息。自协商功能完全由物理层芯片设计实现，因此并不使用专用数据报文或带来任何高层协议开销。

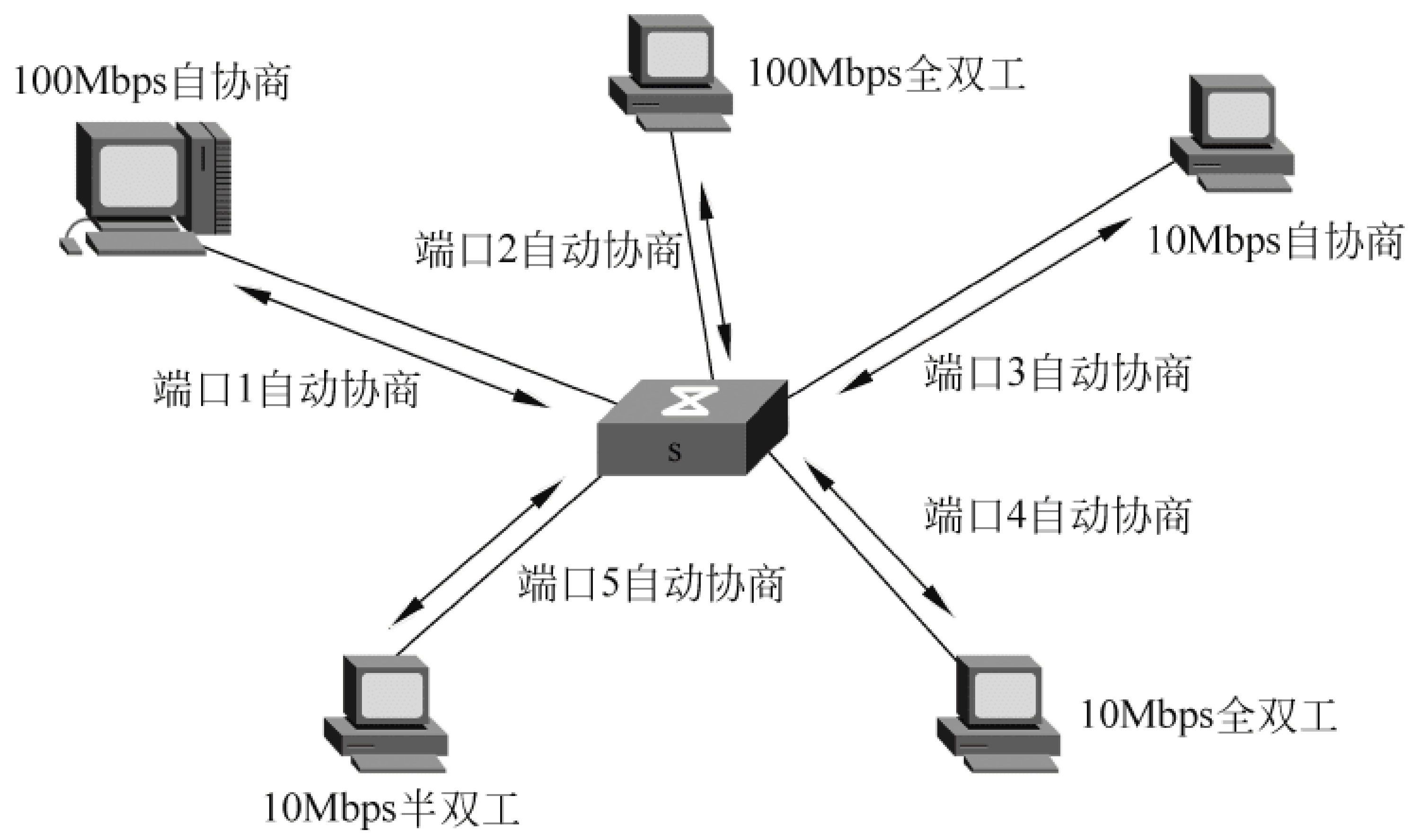


图 7.11 以太网自协商技术的应用

自协商功能的基本机制就是将协商信息封装进一连串的快速连接测试脉冲。每个网络设备必须能够在上电、管理命令发出，或是用户干预时发出此串脉冲，收端将这些信息提取出来就可以得到对端设备支持的工作模式，以及一些用于协商握手机制的其他信息。为了保持与原有 10BASE-T 不具备自协商功能设备的兼容性，自协商协议还具有接受与它们兼容的连接整合性测试脉冲——普通连接脉冲(Normal Link Pulse，NLP)的功能，当一个设备不能对快速连接脉冲做出有效的反应，而仅返回了一个普通连接脉冲时，它将被作为一个 10BASE-T 兼容设备对待。

在链路初始化时，自协商协议向对端设备发送 16 位的报文，并从对端设备接收类似的

报文。根据需要，一个报文可以使用多个16位的“页”，但最常见的协商只需要基本页的操作，如图7.12所示。自协商的内容主要包括速度、全双工、流控等，一方面通知对端设备自身可工作的方式；另一方面，从对端发来的报文中获得对端设备可以工作的方式。

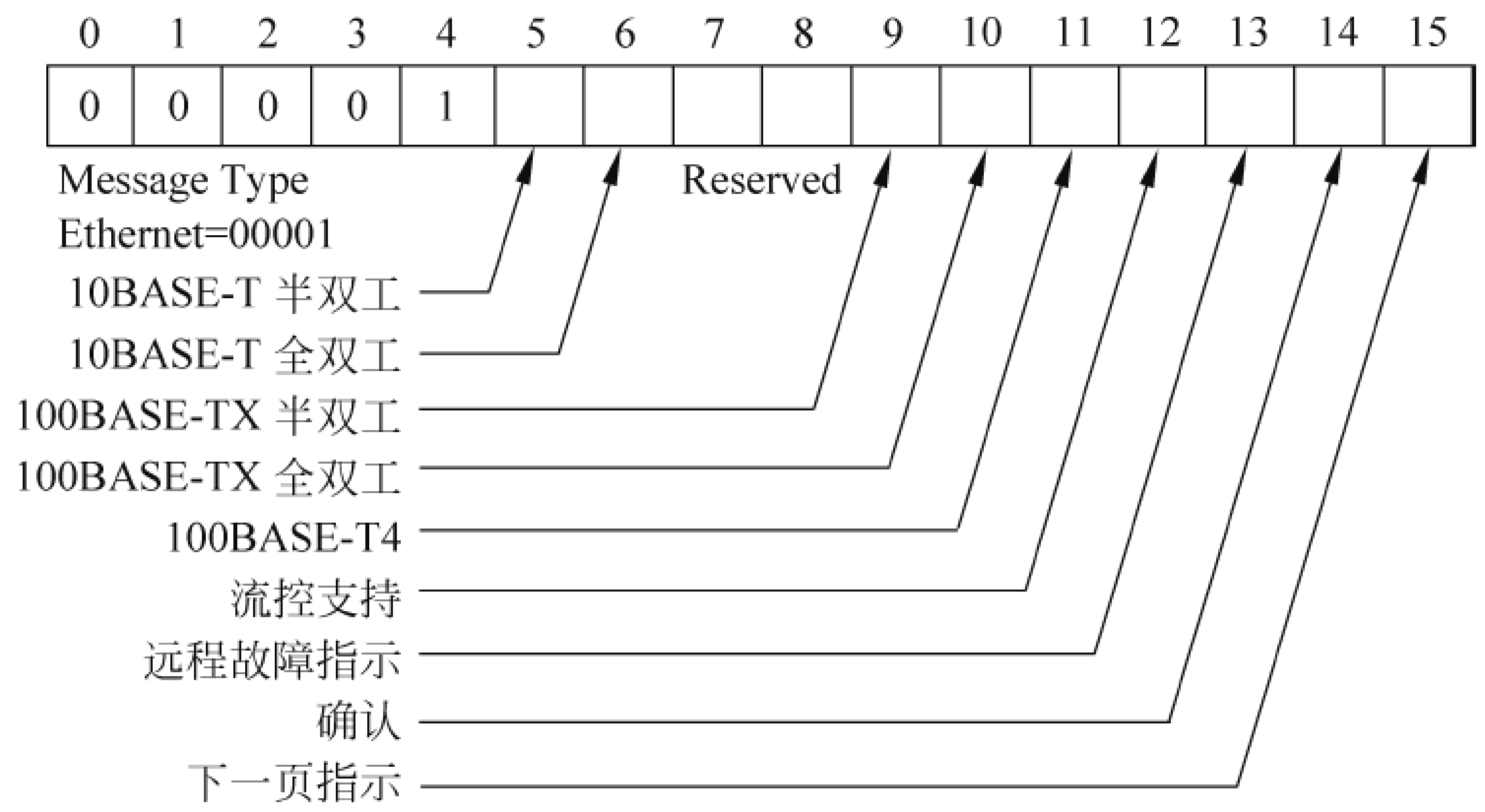


图7.12 以太网自协商技术的16位报文格式

自协商报文的每位之前都插入一个时钟脉冲，是以一系列时钟和数据脉冲的形式发送的，整个报文按16ms间隔重复，指导自协商过程完成。对端设备如不具备自协商功能时，当连接10BASE-T时，是通过识别10BASE-T设备每隔16ms发出的NLP信号来识别对端的设备类型；当连接100BASE-T时，是通过识别信号电平、时序及编码获得连接信息来识别对端的设备类型。所以不通过自协商，同样也可以完成识别10Mbps或100Mbps的功能。

如自协商设备的对端设备不支持自协商，假设它的默认设置为：链路工作于半双工模式，而对端设备为强制10Mbps全双工工作模式，这样的结果将是对端工作在10Mbps全双工工作模式，自协商设备工作在10Mbps半双工的工作模式，这种连接虽然可以通信，但会产生大量的冲突，需要在组网中注意避免。

当协商双方都支持一种以上的工作方式时，需要有一个优先级方案来确定一个最终工作方式。按优先级从高到低（A→E）的顺序列出的IEEE 802.3所支持的5种模式：A（100BASE-TX全双工）→B（100BASE-T4）→C（100BASE-TX）→D（10BASE-T全双工）→E（10BASE-T）。

通常光纤连接很难解决与原有系统的兼容问题，无法很好地实现光纤上的以太网自协商。因此，光纤链路两侧的工作模式通常使用手工配置（速度、双工模式、流量等），如果配置不一样是无法通信的。目前吉比特以太网的自协商机制已基本实现。

2. 智能MDI/MDIX识别技术

如图7.13所示为网络连接示意，采用智能MDI/MDIX（交叉连接/普通直连），就不需要知道电缆另一端为MDI还是MDIX设备，消除了由于电缆配错引起的连接错误，简化了安装维护过程，两种电缆（交叉、直连）的连接可适用于交换机的网络设备。物理层芯片内部的电子开关可以进行MDI和MDIX之间的智能切换。直连网线也称平行网线。

以太网交换机属于MDIX设备，输出的以太网口属于MDIX接口，连接MDI类设备（如PC机）时，需要使用普通直连网线，如果采用交叉网线，是不能正确连接通信的。当前某些最新的以太网交换机，如某些以太网交换机的10/100M以太网口具备智能MDI/MDIX

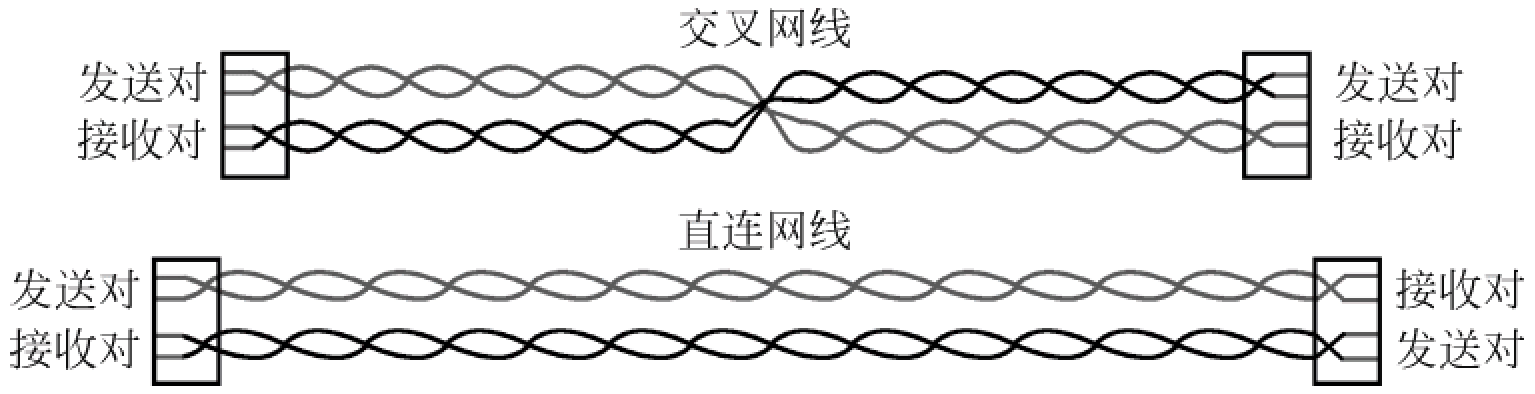


图 7.13　网络连接示意

识别技术，可以自动识别连接的网线类型，用户不管采用普通网线或者交叉网线均可以正确连接设备，极大方便了用户的使用。用户也可以对接口进行配置，将其强制配置成 MDIX 或者 MDI 工作方式。

3．流量控制(PAUSE)

网络拥塞一般是由于线速不匹配(如 100Mbps 向 10Mbps 接口发送数据)和突发的集中传输而产生的，它可能导致的情况有：延时增加、丢包、重传增加、网络资源不能有效利用。IEEE 802.3x 规定了一种 64 字节的"PAUSE"MAC 控制帧的格式。当接口发生阻塞时，交换机向信息源发送"PAUSE"帧，告诉信息源暂停一段时间再发送信息。

PAUSE 功能可以用来控制下列设备之间的数据流：一对终端(简单的两点网络)、一个交换机和一个终端、交换机和交换机之间的链路。PAUSE 功能的增加，是为了防止当瞬时流量过载导致的缓冲区溢出而造成以太网帧的丢弃。假设一个设备用来处理网络上稳定状态的数据传输，并允许随时间变化有一定数量的流量过载，PAUSE 功能可以使这样的设备在流量增长暂时超过其设计水平时，不会发生丢帧现象。该设备通过向对端设备发送 PAUSE 帧，来防止自己内部的缓冲区溢出，而对端设备在接收到 PAUSE 帧后，就会暂时停止发送数据。这样，使第一个设备有时间来减少自己的缓冲拥塞。

IEEE 制定了全双工流量控制标准 802.3x。802.3x 是在双全工环境中去实现流量控制，交换机产生一个 PAUSE 帧，PAUSE 帧使用一个保留的组网地址 01-80-C2-00-00-01，将它发送给正在发送的站，发送站接收到该帧后，就暂停或停止发送。由于 PAUSE 帧使用的是保留组播地址，所以不会被交换机等设备转发，也就不会产生多余的信息量。

PAUSE 不能解决端到端的流量控制问题，不能协调在多个链路上的操作，不能解决持续性流量过载的问题等。

4．接口聚合

接口聚合(pork aggregating)，又称接口捆绑、接口聚集或链路聚集，也属于接口干路。它通过两个网络设备之间的多个接口并行连接，以提高传输带宽。要聚合两端更多的接口，可能会使用更多的扩展模块，可能需要向制造厂家申请最新的驱动软件。用户需要对其进行正确的配置和维护，接口聚合只能工作在全双工模式下，只适用于 802.3 协议族的 MAC 机制，所有捆绑接口的速率必须一致。

1）接口聚合的优点

(1) 增加网络带宽。接口聚合将多个连接的接口捆绑成为一个逻辑连接，捆绑后的带宽是每个独立接口的带宽总和。采用支持该特性的交换机可以增加网络的带宽，如可以将 6 个 100Mbps 接口捆绑在一起组成一个 600Mbps 的连接。

(2) 提高网络连接的可靠性。为了保证高速服务器以及主干网络连接的可靠性，采用接口聚合是一个良好的设计。例如，将一根线缆断开不会导致链路中断，也就是说，组成接

口聚合的某一接口连接失败,网络数据将自动重定向到那些好的连接上。

(3) 接口聚合也可以采用已有的硬件获得,一个支持接口聚合的多接口 100Mbps 的以太网交换机,就可以获得更高的带宽。同样,支持高速服务器的多接口网卡,当上行 1000Mbps 的带宽不够时,将两个 1000Mbps 的接口进行捆绑,获得两倍的上行带宽。

2) 接口聚合的应用

图 7.14 显示了接口聚合的典型应用,它们是交换机之间的连接、交换机到高速服务器或路由器的连接以及高速服务器(或高速路由器)之间的连接。

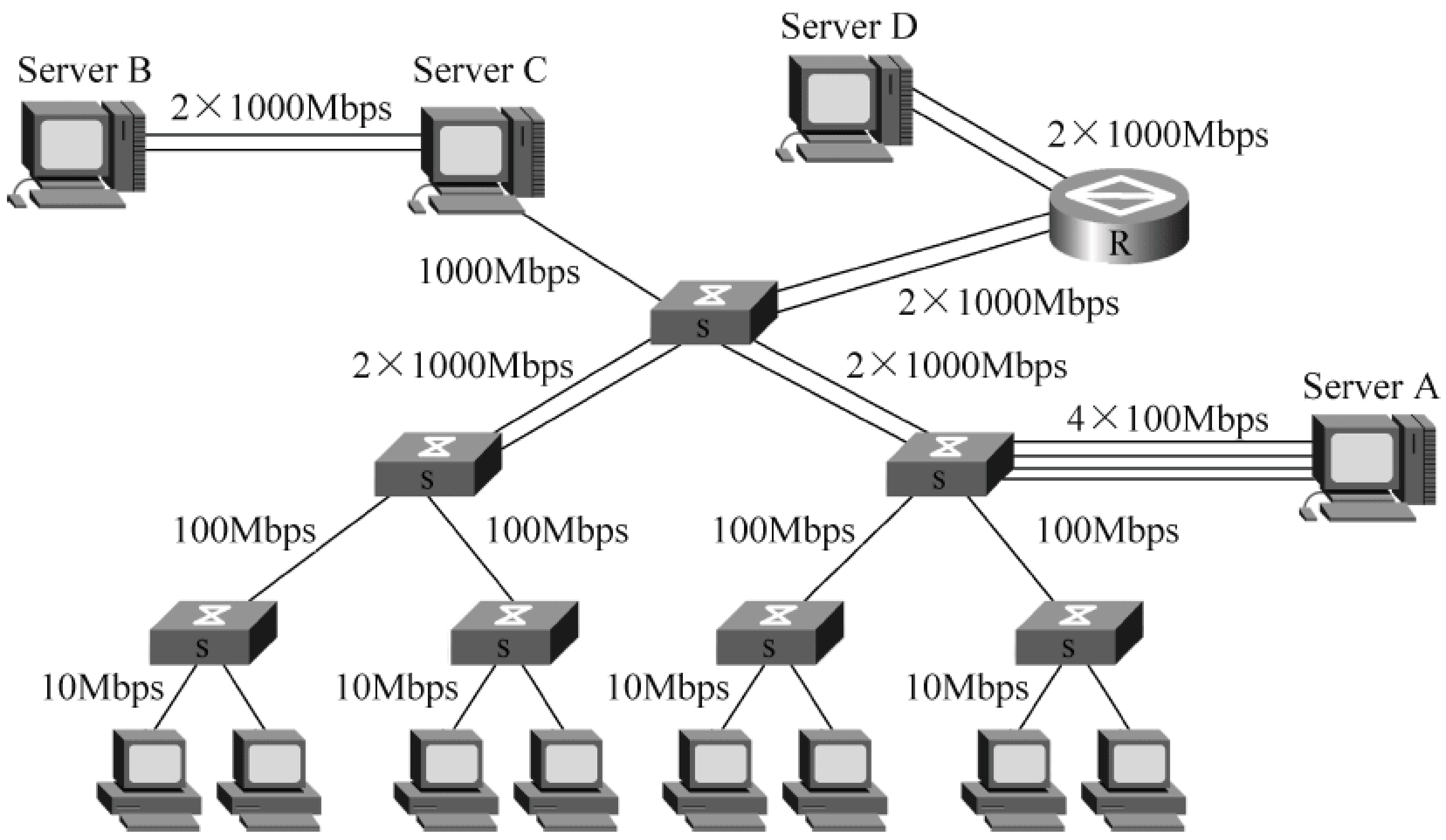


图 7.14 接口聚合应用

(1) 交换机到交换机的连接。

图 7.14 中的两台工作组级交换机之间的连接采用了两个 100Mbps 的接口捆绑成 200Mbps,网络带宽得到了增加,网络连接的可靠性得到了加强,一旦出现某条物理连接故障,网络的带宽变小了,但是网络不会中断,可以保证网络通信不中断。

图中的工作组级交换机是 24 接口 100Mbps 以太网交换机,没有上行高速接口的机型,通过接口聚合技术,无须升级硬件就可以扩展网络带宽,当然,这需要交换机(硬件和软件)支持接口聚合特性。

(2) 交换机到高速服务器或路由器的连接。

图 7.14 中,服务器 A 采用了 4 条 100Mbps 的链路连接到交换机。许多大型服务器具备多个 100Mbps 的网卡,当服务器的访问量增大时,可考虑将多个网卡捆绑成具有更高带宽的接口,满足带宽不断增长的需求。如今,也有一些网卡供应商可以提供将多个 100Mbps 以太网口集中在一块网卡上的产品,其驱动程序可以配置成多个网口的捆绑模式或者单个网口的工作方式,非常适合高速服务器的应用场合。

图中也给出了以太网交换机与高性能路由器的两个 1000Mbps 的接口进行接口捆绑连接的实例。它的应用和连接服务器是相同的。

(3) 高速服务器(或高速路由器)之间的连接。高速终端设备,如服务器、路由器之间的连接也可以应用接口聚合技术。图 7.14 中给出了两台服务器通过 2 条 100Mbps 的接口进行连接的实例。这种高速的连接对于多处理器服务器系统是非常适用的。

5. VLAN 接口类型

交换机的 VLAN 设置中接口有 3 种类型:access、trunk 和 hybrid。其中 hybrid 混合

了 access 和 trunk 的作用，有的交换机 VLAN 接口只有 access 和 trunk 两种类型，有关名词解释如下。

（1）干线（trunk）：指交换机之间的连接，一个 trunk 口上可以同时传送不同 VLAN 的数据包。

（2）接入（access）：指用于连接 PC 的接口。一个 access 口只属于 1 个 VLAN。

（3）混合（hybrid）：用于 PC 之间的连接或交换机与服务器的连接。一个 hybrid 口上也可以同时传送多个 VLAN 的数据包。

（4）标记（tag）：用于指明数据包属于哪个 VLAN，即 VLAN 的 id。若某一接口在 VLAN 设定中被指定为标记接口（tagged port），所有从此接口转发出的数据包上都将有标记（tagged）。若有非标记的数据包进入交换机，则其经过标记接口时，标记将被加上。此时，其将使用在接口上的 pvid 设定作为增加的标记中的 VLAN id 号。

（5）非标记（untag）：没有 VLAN 标记，也就是说所指数据包不属于任何 VLAN。若某一接口在 VLAN 设定中被指定为非标记接口（untagged port），所有从此接口转发出的数据包上都没有标记。若有标记的数据包进入交换机，则其经过非标记接口时，标记将被去除。因为有的网络设备并不支持标记数据包，无法识别标记数据包，因此，需要将与其连接的接口设定为非标记。802.1q 设计的时候为了兼容与不支持 VLAN 的交换机混合部署，特地设计成可以只有一个 VLAN 允许不 tag。

（6）pvid（接口 VLAN id）：是指应对非标记接口的 VLAN id 设定。当非标记数据包进入交换机，交换机将检查 VLAN 设定并决定如何进行转发。例如：一个 IP 包进入交换机接口的时候，如果没有带 tag 头，且该接口上已配置了 pvid，那么，该数据包就会被打上相应的 tag 头；如果进入的 IP 包已经带有 tag 头，属于 VLAN 数据，那么交换机就不会再增加 tag 头号。

通过表 7.2 给出的 VLAN 各接口收发数据描述，可以更好地理解各类接口的运行功能。

表 7.2　各接口对收发数据的描述

接口	报文	描　述
access	收	收到 PC 机报文，判断是否有 VLAN 信息。如果没有，则打上接口的 pvid，并进行交换转发；如果有，则丢弃
	发	将报文的 VLAN 信息剥离，直接发送到 PC 机
trunk	收	收到一个交换机的报文，判断是否有 VLAN 信息，如果没有，则打上接口的 pvid，并进行转发；如果有，则判断该接口是否允许该 VLAN 的数据进入，只有允许才转发，否则直接丢弃
	发	比较接口的 pvid 和将要发送报文的 VLAN 信息，如果两者相等，则剥离 VLAN 信息后再发送；如果不相等则直接发送
hybrid	收	收到一个 PC 机或交换机的报文，判断是否有 VLAN 信息，如果没有，则打上接口的 pvid，并进行转发；如果有，则判断该接口是否允许该 VLAN 的数据进入；如果可以则转发，否则丢弃
	发	判断该 VLAN 在本接口的属性，如果是 untag，则剥离 VLAN 信息，再发送；如果是 tag，则直接发送。用 disp interface 命令可以看到该接口对哪些 VLAN 是 untag，哪些 VLAN 是 tag

7.3.3 以太网接口配置

在前面介绍交换机有关以太接口时,给出的都是 Cisco 命令,这里以华为设备命令为例,介绍在以太网接口状态视图下进行的基本设置。一般情况下如果要取消或删除前面的配置,在原配置命令前面加 undo 即可,所以 undo 命令就不单独介绍。

1. 配置接口的工作速率

```
speed { 10 | 100 | auto }                    !设置以太网接口的速率
```

默认情况下,接口的速率处于 auto(自协商)状态。以太网接口支持 10Mbps、100Mbps 两种速率,而吉比特以太网接口只能支持 1000Mbps 速率,不能设置。

2. 配置接口的全双工操作

```
duplex { half | full | auto }                !设置以太网接口的双工状态
```

通过 duplex 命令,可以对以太网接口的双工特性进行设置,如全双工(full)、半双工(half)或自协商(auto)状态。默认情况下,接口的双工状态为 auto 状态。

【例 4】 将接口 0/3 强制设置为 100M 全双工工作状态。

答:

```
[S3526]interface ethernet 0/3
[S3526 - Ethernet0/3]speed 100
[S3526 - Ethernet0/3]duplex full
```

3. 配置接口的流控

```
flow - control                               !开启以太网接口的流量控制
```

默认情况下,接口的流量控制为关闭状态(disable)。

4. 配置接口的 MDI/MDIX 工作方式

```
mdi { normal | across | auto }               !设置以太网接口连接的网线的类型。
```

网线类型为普通(normal)、交叉(across)、自动识别(auto),默认情况下为 auto 型。

5. 配置接口聚合(干路)

接口聚合,也称接口汇聚,是将多个接口聚合在一起,以实现对出/入负荷在各成员接口中进行分担。如 S3526 交换机中只有 3 个以太网负荷分担组和一个光口负荷分担组,要求组内的接口必须连续。

```
link - aggregation port_num1 to port_num2 { ingress | both }     !设置以太网汇聚接口
```

该命令用来设置以太网接口汇聚形成的汇聚接口,命令中参数含义如下。

port_num1 表示加入的以太网物理接口范围 1。

port_num2 表示加入的以太网物理接口范围 2。

ingress 表示根据源 MAC 地址进行数据帧的分发。

both 表示根据源 MAC 地址和目的 MAC 地址进行数据帧的分发。

6. 查看当前接口干路配置

```
display link-aggregation [master_port_num]      !显示汇聚接口的信息
```

汇聚接口信息包括汇聚接口等。如果不指定 master_port，则显示所有的干线接口。需要注意的是物理连接上的两端设备均要设置。

【例 5】 将接口 1～2 配置为干路视图。

答：

```
[S3526]link-aggregation ethernet0/1 to ethernet0/2 ingress      !设置以太网汇聚接口 1～2
[S3526]display link-aggregation ethernet0/1      !显示汇聚接口 ethernet0/1 的信息
Master port: Ethernet0/1                         !显示 Ethernet0/1 为主接口
Other sub-ports: Ethernet0/2                     !显示 Ethernet0/2 为子接口
Mode: ingress                                    !显示的是根据源 MAC 地址进行数据帧的分发
```

7. 显示接口的配置信息

接口信息包括接口类型、接口状态、是否双工、接口速率、流控、广播抑制比、是否汇聚接口等，下面给出的是 display 命令。

```
display interface[port_num]   !显示接口的所有信息。port_num 为单个以太网接口，即
                              !port_num = {interface_type interface_number | interface_name}
```

【例 6】 显示接口配置 ethernet 0/2 信息。

答：

```
[S3526]display interface ethernet 0/2
```

最后为了学习方便，有关显示的以太网接口配置信息项注释由表 7.3 具体列出。

表 7.3　接口配置信息描述

信　息　项	注　　释
Ethernet2/1/1 current state	以太网接口当前处于开启或关闭状态
IP Sending Frames' Format	以太网帧格式
Hardware address	接口硬件地址
The Maximum Transmit Unit	最大传输单元
Media type	介质类型
loopback not set	接口环回测试状态
Port hardware type	接口硬件类型
100Mbps-speed mode, full-duplex mode Link speed type is auto-negotiation, link duplex type is auto-negotiation	接口的双工状态和速率均设置为自协商状态，与对端协商的实际结果是 100Mbps 速率和全双工模式
Flow-control is not supported	接口流控状态
The Maximum Frame Length	接口允许通过的最大以太网帧长度
Broadcast MAX-value	接口广播抑制带宽
Allow jumbo frame to pass	接口允许长帧通过
Port VPN status	接口 VPN 状态
PVID	接口默认为 VLAN ID
Mdi type	网线类型

续表

信　息　项	注　　释
Port link-type	接口链路类型
Tagged VLAN ID	标识在该接口有哪些 VLAN 的报文需要打 tag 标记
Untagged VLAN ID	标识在该接口有哪些 VLAN 的报文不需要打 tag 标记
Last 300 seconds input：0 packets/sec 0 bits/sec Last 300 seconds output：0 packets/sec 0 bits/sec	接口最近 300s 输入和输出速率

习题

1. HDLC、PPP 协议是运行在 OSI 七层参考模型哪一层的协议？分别简述其基本原理。

2. PPP 的简单配置可以总结为两步。第一步，配置接口的链路层协议为 PPP：link-protocol ppp；第二步，配置 PPP 验证方式及用户名、用户口令。PPP 有 PAP 验证和 CHAP 验证，要求写出用于验证、被验证方的有关命令。

3. 根据图 7.15，完成以下配置。本题最好能在实验环境下完成。

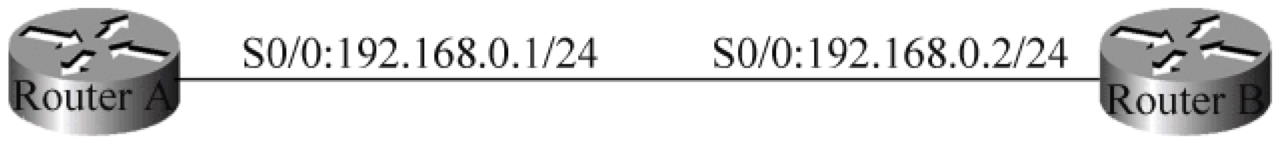


图 7.15　HDLC、PPP 配置环境

(1) 配置各路由器的 IP 地址等参数。配置路由器 A 的 serial0/0 时钟频率为 64 000。

(2) 配置路由器 A、B 间的串行接口，封装类型为 HDLC。

(3) 改变路由器 A 和路由器 B 串行接口 serial0/0 的封装格式为 PPP。

(4) 配置路由器 A 和路由器 B 的串行接口 serial0/0，进行 PAP 认证。

(5) 配置路由器 A 为 PAP 认证服务器，建立本地用户名为 routera，口令 123。

(6) 配置路由器 B 为 PAP 认证客户端，将用户名 routerb 和口令 123 发送到对端路由器。

4. 举例说明常用的以太网设置命令。

第8章 数据网络配置

随着接入 Internet 用户数量的猛增，IP 地址（特别是 IPv4）资源也就愈加显得捉襟见肘，网络地址转换（NAT）能够有效地将私有地址转化为合法 IP 地址，它不仅解决了 IP 地址不足的问题，还有效地避免了来自网络外部的攻击，被广泛应用于 Internet 接入网络中；访问控制列表（ACL）可以用来控制各类网络设备接口的进出数据报文，适用于所有的路由协议，是过滤网络、控制流量、提升网络安全的一种技术手段。本章重点介绍 ACL、NAT，以及 X.25、帧中继等有关网络配置，以使读者掌握有关数据网络协议在路由、交换设备中的应用。

8.1 访问控制列表

8.1.1 ACL 概述

1. ACL 的分类及特性

访问控制列表（Access Control List，ACL）应用在路由器的接口上，通过匹配数据包信息与访问表参数来决定允许（permit）数据包通过，还是拒绝（deny）数据包通过某个接口。

访问列表类型可分为标准 IP 访问列表和扩展 IP 访问列表。

标准访问列表：其只检查数据包的源地址，从而允许或拒绝基于网络、子网或主机的 IP 地址的所有通信流量通过路由器的出口。

扩展 IP 访问列表：它不仅检查数据包的源地址，还要检查数据包的目的地址、特定协议类型、源端口号、目的端口号等。

标准的 IP 访问列表只匹配源地址，一般都使用扩展的 IP 访问列表以达到精确的要求。标准的访问列表尽量靠近目的，由于标准访问列表只使用源地址，因此将其靠近源，会阻止报文流向其他端口。扩展的访问列表尽量靠近过滤源的位置上，以免访问列表影响其他接口上的数据流。

ACL 的相关特性：路由器的每一个接口可以在进入（inbound）和离开（outbound）两个方向上分别应用一个 ACL，且每个方向上只能应用一个 ACL。在路由选择进行以前，应用在接口进入方向的 ACL 起作用。在路由选择决定以后，应用在接口离开方向的 ACL 起作用。

2. 路由器对访问控制列表的处理过程

（1）如果接口上没有 ACL，就对这个数据包继续进行常规处理。

（2）如果对接口应用了访问控制列表，与该接口相关的一系列 ACL 语句组合进行检查：

① 若第一条不匹配，则依次往下走，直到有一条语句匹配，则不再继续判断，路由器将

决定该数据包允许通过或拒绝通过。

② 若进行到最后也没有任一语句匹配，则路由器根据默认处理方式丢弃该数据包。

③ 基于 ACL 的测试条件，数据包不是被允许，就是被拒绝。

(3) ACL 的出与入。使用命令 ip access-group 可以把访问控制列表应用到某一个接口上。通过 in 或 out 指明访问控制列表是对进来的，或是对出去的数据包进行控制。

8.1.2 ACL 配置

ACL 配置步骤：首先是配置访问列表语句，然后将配置好的 ACL 应用到某个接口上。

1. 标准 IP 访问列表的配置

(1) 标准访问列表。

```
access - list access - list - number deny|permit source - address source - wildcard [log]
```

access-list-number：表示列表号，只能是 1～99 之间的一个数字。

deny|permit：deny 表示匹配的数据包将被过滤掉；permit 表示允许匹配的数据包通过。

source-address：表示单台或一个网段内的主机的 IP 地址。

source-wildcard：通配符掩码，通配符掩码设置为 1 时，IP 地址的对应位可以是 0 或 1；如果设为 0 时，IP 地址的对应位必须被精确匹配。例如通配符掩码为 0.0.255.255 时，表示 IP 地址的前 16 位必须精确匹配。

通配符掩码的两种特殊形式：一个是 host，表示一种精确匹配，是通配符掩码 0.0.0.0 的简写形式，句子将 host 放在源地址的前面，如 196.168.1.2 0.0.0.0 表示为 host 196.168.1.2；另一个是 any，表示全部不进行匹配，是通配符掩码 255.255.255.255 的简写形式，如 196.168.1.2 255.255.255.255 表示为 any。

log：访问列表日志，如果该关键字用于访问列表中，则对匹配访问列表中条件的报文作日志。

(2) 配置标准 IP 访问列表。

标准 IP 访问列表的配置命令：

```
ip access - group access - list - number in|out
```

in：通过接口进入路由器的报文。

out：通过接口离开路由器的报文。

(3) 显示所有协议的访问列表配置细节。

```
show access - list [access - list - number]
```

(4) 显示 IP 访问列表。

```
show ip access - list [access - list - number]
```

【例 1】 各个主机经过 RouterB 的 FastEthernet 1/0 连接到局域网上，要限制其中一台主机 hostA(192.168.1.2)访问服务器 server，而该局域网的其他主机可以访问服务器 server。

答：标准 IP 访问列表的配置举例如下。

```
RouterB# configure terminal
RouterB(config)# access - list 1 deny host 192.168.1.2     !定义列表 1 拒绝 hostA 访问
RouterB(config)# access - list 1 permit any                !定义列表 1 允许其他主机访问
RouterB(config)# interface FastEthernet 1/0          !在连接局域网的 f1/0 接口上启用列表 1
RouterB(config - if)# ip access - group 1 in         !在进入 RouterB 接口 f1/0 时,一律执行列表 1
RouterB(config - if)# end
RouterB# Show ip access - list 1                           !显示访问控制列表 1 有关信息
```

2. 扩展 IP 访问列表的配置

配置扩展访问列表命令如下。

```
access - list [list number] [permit | deny] [protocol] [source address source - wildcard]
[source port] [destination address destination - wildcard] [destination port]
operator port [port] [established]
```

access-list-number：编号范围为 100～199。

permit|deny：表示在满足条件的情况下，该入口是允许还是拒绝后面匹配的地址的通信。

protocol：需要被过滤的协议的类型，如 ip、tcp、udp、icmp、eigrp、gre、ospf 等。

source-address：源 IP 地址。

source-wildcard：源通配符掩码。

source-port：源端口号，可以是单一的某个端口，也可以是一个端口范围。

destination-address：目的 IP 地址。

destination-wildcard：目的地址通配符掩码。

destination-port：目的端口号，指定方法与源端口号的指定方法相同。

operator port [port]：定义端口规则，用于测试数据包的端口号是否匹配。operator 是运算符，有 eq(等于)、neq(不等于)、lt(小于)、gt(大于)、range(范围)几种。port [port]是端口号，如使用 range 算符时，需指定两个端口号，其他都只指定一个；如为 eq port-name 时，用于测试数据包的端口号是否匹配。其中 eq 是等于算符，port-name 是端口名称。端口名称如下：

TCP 端口名称有 bgp、pop3、smtp、telnet、www、ftp 等；

UDP 端口名称有 domain、echo、mobile-ip、rip、snmp、tftp 等。

established：如果数据包使用一个已建立连接，例如该数据包设置了 ACK(确认序号有效)位，便可以允许 TCP 信息量通过。

【例 2】　要求允许 LAN3(192.168.3.0/24)的所有主机能登录 Internet，但只能浏览 WWW、FTP、SMTP、POP3 协议的通信。LAN2(192.168.2.0/24)中的主机 192.168.2.2 向 Internet 提供 FTP 服务，主机 192.168.2.3 向 Internet 提供 SMTP 服务，主机 192.168.2.4 向 Internet 提供 WWW 服务，其余主机则不能被 Internet 访问。LAN1(192.168.1.0/24)中的主机不能访问 Internet，但可以访问 LAN2 和 LAN3。有关路由器及接口可随意指定。

答：

```
Router#configure terminal
Router(config)#access-list 111 permit tcp 192.168.3.0 0.0.0.255 any eq www
!定义 access-list 111: 允许 192.168.3.0/24(TCP 传输)所有主机能访问 WWW 服务器
Router(config)#access-list 111 permit tcp 192.168.3.0 0.0.0.255 any eq ftp
!定义 access-list 111: 允许 LAN3(192.168.3.0/24)的所有主机能访问 FTP 服务器
Router(config)#access-list 111 permit tcp 192.168.3.0 0.0.0.255 any eq smtp
!定义 access-list 111: 允许 LAN3(192.168.3.0/24)的所有主机能访问 SMTP 服务器
Router(config)#access-list 111 permit tcp 192.168.3.0 0.0.0.255 any eq pop3
!定义 access-list 111: 允许 LAN3(192.168.3.0/24)的所有主机能访问 POP3 服务器
Router(config)#access-list 111 deny ip any 192.168.1.0 0.0.0.255
!定义 access-list 111: 拒绝 LAN1(192.168.1.0/24)主机访问服务器
Router(config)#access-list 112 permit tcp any host 192.168.2.2 eq ftp !
!定义 access-list 112: 允许 Internet 任何主机访问 FTP 服务器(IP 为 192.168.2.2)
Router(config)#access-list 112 permit tcp any host 192.168.2.3 eq smtp
!定义 access-list 112: 允许 Internet 任何主机访问 SMTP 服务器(IP 为 192.168.2.3)
Router(config)#access-list 112 permit tcp any host 192.168.2.4 eq www!
!定义 access-list 112: 允许 Internet 任何主机访问 WWW 服务器(IP 为 192.168.2.4)
Router(config)#interface Serial 1              !路由器通过 s1 连接 LAN1、LAN3
Router(config-if)#ip access-group 111 out      !通过 s1 出方向执行 access-list 111
Router(config-if)#exit
Router(config)#interface FastEthernet 1/0      !路由器通过 f1/0 连接 LAN2
Router(config-if)#ip access-group 112 out      !通过 f1/0 出方向执行 access-list 112
Router(config-if)#end
Router#show ip access-lists 111                !查看 access-list 111 有关信息
Router#show ip access-lists 112                !查看 access-list 112 有关信息
```

3. 基于名称方式创建的标准或扩展 ACL

IP 访问列表可以是一个编号，也可以通过一个名称来命名，命名能更直观地反映出访问列表完成的功能。命名访问列表突破了 99 个标准访问列表和 100 个扩展访问列表的数目限制，能够定义更多的访问列表。命名 IP 访问列表允许删除个别语句，而编号访问列表只能删除整个访问列表。单个路由器上命名访问列表的名称在所有协议和类型的命名访问列表中必须是唯一的，而不同路由器上的命名访问列表名称可以相同。基于名称方式的 ACL 命令如下：

```
ip access list [standard | extended] [name]
```

standard|extended：指出是基于名称的标准 ACL，还是基于名称的扩展 ACL。前面命令中的过滤规则，可选用的参数选项在这里同样对应适用。

name：名称取代了前面介绍的 access-list-number，用起来更为直观。

用此命令建立了一个 ACL 后，后面所有的 permit|deny 规则就都可以在这个模式下进行配置，如 permit 192.168.2.1 0.0.0.0 就是一条标准的 ACL 过滤语句。

【例 3】 针对前面讲的例 2，要求用命名访问列表配置进行配置。

答：

```
Router#configure terminal
Router(config)#ip access-list extended acl_lan1_lan3   !定义扩展 ACL 名为 acl_lan1_lan3
```

```
Router(config-ext-nacl)#permit tcp 192.168.3.0 0.0.0.255 any eq www
Router(config-ext-nacl)#permit tcp 192.168.3.0 0.0.0.255 any eq ftp
Router(config-ext-nacl)#permit tcp 192.168.3.0 0.0.0.255 any eq smtp
Router(config-ext-nacl)#permit tcp 192.168.3.0 0.0.0.255 any eq pop3
Router(config-ext-nacl)#deny ip any 192.168.1.0 0.0.0.255
Router(config-ext-nacl)#exit
Router(config)#ip access-list extended acl_lan2     !定义扩展 ACL 名为 acl_lan2
Router(config-ext-nacl)#permit tcp any host 192.168.2.1 eq www
Router(config-ext-nacl)#permit tcp any host 192.168.2.2 eq ftp
Router(config-ext-nacl)#permit tcp any host 192.168.2.3 eq smtp
Router(config-ext-nacl)#exit
Router(config)#interface Serial 1
Router(config-if)#ip access-group acl_lan1_lan3 out
Router(config-if)#interface interface FastEthernet 1/0
Router(config-if)#ip access-group acl_lan2 out
Router(config-if)#end
Router#show ip access-lists acl_lan1_lan3
Router#show ip access-lists acl_lan2
```

4. TCP 拦截原理及其配置

1）TCP 拦截模式

我们知道建立 TCP 的 3 次握手：源主机发送一个含 SYN 标志的 TCP 报文给目标主机；目标主机在收到客户端的 SYN 报文后，将返回一个 SYN+ACK 的报文，表示源主机的请求被接收；源主机返回一个 ACK 确认报文给目标主机，这样一个 TCP 连接就得以完成。所谓 TCP 拦截就是防止 SYN 泛洪攻击，就是在 TCP 连接请求到达目标主机之前，通过拦截和验证来阻止网络受到的攻击，也就是说 TCP 拦截可以在拦截和监视两种模式下工作。其主要作用如下。

（1）在拦截模式下，路由器拦截 TCP 到达的 SYN 请求，并代表目标主机建立与源主机的连接，如果连接成功，则源主机建立与目标主机的连接，并将两个连接进行透明合并。在整个连接期间，路由器会一直拦截和发送数据包。对于非法的连接请求，路由器会提供更为严格的半连接超时限制，以防止自身的资源被 SYN 攻击耗尽。

（2）在监视模式下，路由器被动地观察流经路由器的连接请求，如果连接超过了所配置的建立时间，路由器就会关闭此连接。

2）IP 地址欺骗对策

入站方向：不允许任何源地址是内部主机或内部网络地址的数据包进入一个私有网络。

出站方向：不允许任何源地址不是内部主机或内部网络地址的数据包出站。

3）TCP 拦截配置举例

【例 4】 DNS 服务器的 IP 地址为 161.86.7.10，要求配置路由器进行 TCP 拦截，以保护 DNS 服务器不受 SYN 泛洪攻击。要求采用拦截模式，设置最大半连接数的高、低值分别为 400 和 200；每分钟保持连接数的高、低值分别为 500 和 300。

```
Router#configure terminal
Router(config)#access-list 111 permit tcp any host 161.86.7.10 !设置允许 DNS 的 TCP 连接
Router(config)#ip tcp intercept mode intercept                !采用拦截模式
Router(config)#ip tcp intercept max-incomplete high 400       !最大半连接数高值
Router(config)#ip tcp intercept max-incomplete low 300        !最大半连接数低值
```

```
Router(config)# ip tcp intercept one-minute high 500        !每分钟保持连接数的高值
Router(config)# ip tcp intercept one-minute low 300         !每分钟保持连接数的低值
Router(config)# ip tcp intercept list 111                   !开启 TCP 拦截
```

8.2 网络地址转换

8.2.1 NAT 简述

1. NAT 技术

网络地址转换(Network Address Translation,NAT)是一种将私有(保留)地址转化为合法 IP 地址的转换技术。借助于 NAT,私有地址所在的内部网络通过路由器发送数据包时,私有地址将被转换成合法的 IP 地址,一个局域网只需使用少量 IP 地址即可实现私有地址网络内所有计算机与 Internet 的通信需求。NAT 将自动修改 IP 报文的源 IP 地址和目的 IP 地址,有些应用程序将源 IP 地址嵌入到 IP 报文的数据部分中,所以还需要同时对报文进行修改,以匹配 IP 头中已经修改过的源 IP 地址。虽然 NAT 可以借助于某些代理服务器来实现,但很多时候都是在路由器上实现的。

2. 内部网络地址和外部网络地址

NAT 起到将内部私有地址翻译成外部合法的全局地址的功能,它使得不具有合法 IP 地址的用户可以通过 NAT 访问到外部 Internet。当建立内部网时,可使用私有网络地址用于主机,内部网络使用的私有网络的地址为:

1 个 A 类地址 10.0.0.0/8;

16 个 B 类地址 172.16.0.0/16～172.31.0.0/16;

256 个 C 类地址 192.168.0.0/16。

NAT 配置中可能会用到以下地址,注意区别。

内部本地地址(inside local address):内部网络主机使用的私有 IP 地址。

内部全局地址(inside global address):内部网络使用的公有 IP 地址。

外部本地地址(outside local address):外部网络主机使用的本地公有 IP 地址。

外部全局地址(outside global address):外部网络主机使用的公有 IP 地址。

8.2.2 NAT 配置命令

以下给出有关 NAT 的命令。

1. 访问控制列表命令(在前面已讲过)

```
access-list access-list-number {deny|permit|remark} {source [source-wildcard]|any}
```

【例 5】 将 172.16.1.1～172.16.254.254 定义为允许访问,访问列表号为 1。

答:

```
RouterA(config)# access-list 1 permit 172.16.0.0 0.0.255.255
```

2. 配置接口的类型

```
ip nat {inside | outside}
```

inside：表示该接口连接内部网络。

outside：表示该接口连接外部网络。

此命令是在接口配置模式下使用，用于指定 NAT 的内网和外网的接口。数据包只有在 outside 接口和 inside 接口之间路由时，并且符合一定规则，才会进行 NAT 转换。所以实现 NAT 的路由器必须配置至少一个 outside 接口和一个 inside 接口，也可配置多个。使用 no ip nat 命令后，可使接口不再应用 NAT。

【例 6】 要求配置路由器 A，FastEthernet 0/0 连接的是内网，定义为 inside 接口；Serial 2/0 连接的是外网，定义为 outside 接口。

答：

```
RouterA(config)# interface FastEthernet 0/0
RouterA(config-if)# ip address 192.168.1.1 255.255.255.0
RouterA(config-if)# ip nat inside                      !f0/0 确定为内部接口
RouterA(config-if)# no shutdown
RouterA(config-if)# exit
RouterA(config)# interface Serial 2/0
RouterA(config-if)# ip address 210.19.12.17 255.255.255.0
RouterA(config-if)# ip nat outside                     !Serial 2/0 确定为连接外网接口
RouterA(config-if)# no shutdown
RouterA(config-if)# exit
RouterA(config)#
```

3. 配置内部全局地址池命令

ip nat pool name start-ip end-ip {**netmask** netmask | **prefix-length** prefix-length} [type rotary]

此命令定义一个全局地址池。其中：

name：地址池名。

start-ip：地址池开始 IP 地址。

end-ip：地址池结束 IP 地址。

netmask：掩码，prefix-length 为掩码长度，也就是网络占用的二进制位数。

type rotary：轮流、均衡占用。

4. NAT 内部目标地址转换命令

ip nat inside destination list access-list-number **pool** pool-name

本命令在全局配置模式下，启用 NAT 内部目标地址转换。NAT 内部目标地址转换可用于实现 TCP 负载均衡，可以用一台虚拟主机代替多台实际主机接收用户的 TCP 请求，由 NAT 把这些请求轮流映射到各个实际主机上，达到负载分流的目的。使用 no 命令为关闭。其中：

access-list-number：访问控制列表的表号。它指定由哪个访问控制列表来定义目标地址的规则。配置 TCP 负载均衡时，访问控制列表定义的是虚拟主机的地址，IP 地址池中定义的是各台实际主机的地址。

pool-name：IP 地址池名字。该地址池定义了用于 NAT 转换的内部本地地址。

默认值：没有启用 NAT 内部目标地址转换。

【例 7】 要求配置路由器 A 定义为一个 TCP 负载均衡，虚拟主机地址为 210.19.12.1，由 access-list 2 定义；实际主机地址为 192.168.1.1～192.168.1.9，由地址池 aaa 定义。

答：

```
RouterA(config)#ip nat pool aaa 192.168.1.1 192.168.1.9 netmask 255.255.255.0 type rotary
RouterA(config)#access-list 2 permit 210.19.12.1 0.0.0.0
RouterA(config)#ip nat inside destination list 2 pool aaa
```

5. 配置内部动态源地址转换命令

ip nat inside source list access-list-number {**pool** pool-name | **interface** interface-type} [**overload**]

access-list-number：访问列表编号。

access-list-number：访问控制列表的表号。它指定由哪个访问控制列表来定义源地址的规则。

pool-name：IP 地址池名字。该地址池定义了用于 NAT 转换的内部全局地址。

interface-type：接口类型，如 f0/0，指定用该接口的 IP 地址作为内部全局地址。

overload：启用接口复用，允许将多个内部本地地址转换为一个内部全局地址。使每个内部全局地址可以和多个内部本地地址建立映射，允许多个内部本地地址使用相同的内部全局地址。如果不加 overload，则一个时间段内只有一条内部全局地址可以通信，不能复用。

此命令是在全局配置模式下，用于建立动态源地址翻译。启用内部源地址转换的动态 NAT。在配置时，访问控制列表定义的是内部本地地址的规则，IP 地址池中定义的是内部全局地址，它通常是注册的合法地址。

【例 8】 要求路由器 A 定义一个内部源地址动态 NAT，内部本地地址为 192.168.1.×的格式，由 access-list 1 定义，允许进行 NAT 转换；内部全局地址为 210.10.10.1～210.10.10.5，由地址池 aaa 定义。每个全局地址都可以和多个本地地址建立映射，用端口号区分各个映射。

答：

```
RouterA(config)#ip nat pool aaa 210.10.10.1 210.10.10.5 netmask 255.255.255.0
RouterA(config)#access-list 1 permit 192.168.1.0 0.0.0.255
RouterA(config)#ip nat inside source list 1 pool aaa overload
```

【例 9】 要求路由器 A 定义一个内部源地址动态 NAT，内部本地地址为 172.16.×.×的格式，由 access-list 2 定义。内部全局地址为路由器串口 Serial 1/0 对应的 IP 地址，所有本地地址都会映射为这一个 IP 地址，用端口号区分各个映射。

答：

```
Router(config)#access-list 2 permit 172.16.0.0 0.0.255.255
Router(config)#ip nat inside source list 2 interface Serial 1/0 overload
```

6. 内部源地址转换的静态 NAT 命令

ip nat inside source static local-address global-address[permit-inside]
ip nat inside source static protocol local-address local-port global-address global-port

local-address：内部本地地址。是主机在网络内部的 IP 地址，一般是私有地址。

global-address：内部全局地址。是内部主机在外部网络表现出的合法地址。

protocol：协议。可以是 TCP 或 UDP。

local-port：本地地址的服务端口号。

global-port：全局地址的服务端口号，它可以和 local-port 相同或不同。

permit-inside：允许内部用户使用全局地址访问本地主机。

【例 10】 路由器 A 连接一个内部主机，内部本地地址为 192.168.1.2；外部连接的本地全局地址为 210.10.10.1，要求建立内部源地址转换的静态 NAT 配置。

答：

```
RouterA(config)＃ip nat inside source static 192.168.1.2 210.10.10.1
```

【例 11】 路由器 A 连接一个内部 FTP 服务器，内部 IP 地址为 192.168.1.2，外部 IP 地址为 210.10.10.1，则需要配置两条 NAT，因为 FTP 有两个端口：20 和 21，即传数据和指令的端口。要求完成静态 NAT 配置。

答：

```
RouterA(config)＃ip nat inside source static tcp 192.168.1.2 21 210.10.10.1 21
RouterA(config)＃ip nat inside source static tcp 192.168.1.2 20 210.10.10.1 20
```

7. 其他命令

查看生效的 NAT 设置：**show ip nat translations**

如：RouterA(config)＃show ip nat translations

查看 NAT 统计信息：**show ip nat statistics**

如：RouterA(config)＃show ip nat statistics

清除动态 NAT 配置：**clear ip nat translation**

如：清除所有动态 NAT 配置：RouterA(config)＃clear ip nat translation

8.2.3 NAT 配置实例

1. NAT 地址转换概述

1）NAT 配置类型

NAT 配置类型分为：静态地址转换(static address translation)，也称静态 NAT(Static NAT)；动态地址转换(dynamic address translation)，也称动态 NAT(Pooled NAT)；网络地址端口转换(network address port translation)，也称 PNAT(Port-Level NAT)等也属于动态 NAT。静态 NAT，也就是一个私有地址对应一个公网地址，主要用在企业对外提供网络服务的服务器上，如网页服务器。而动态 NAT，一般是多个私网地址对应一个公网地址，然后再通过配置传输层对应的端口号来区分私网内不同的 PC 及同一 PC 的不同进程。

2）网络环境

根据图 8.1 所示，RouterA 执行 NAT 协议，内部网络本地 IP 地址段为 192.168.1.0/28，内部网关 IP 地址为 192.168.1.1/28，网关外网 IP 地址为 68.192.160.1。除网关使用了一个公网 IP 外，还有 5 个公网 IP 可以使用(68.192.160.2～68.192.160.6/29)，也就是内部全局地址。另外，公网上有一台 Web 服务器，外部全局地址 IP 为 202.113.64.2，使用 80 端

口发布 Web 服务。

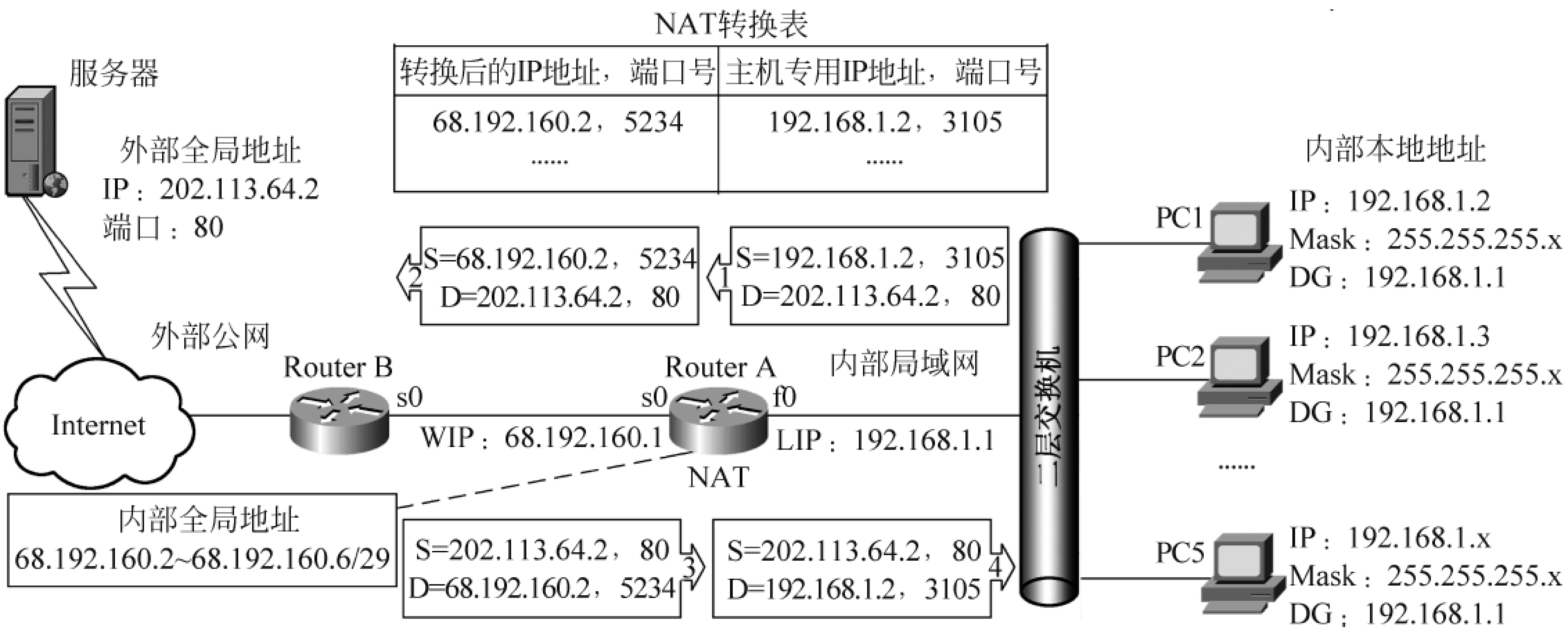


图 8.1　路由器 NAT 转换

3）端口地址转换工作过程

当内部局域网主机需要与 Internet 建立连接时，首先将请求发送到端口 NAT 服务器。NAT 服务器接收到请求后，根据接收到的数据包，检查端口 NAT 映射表。如果还没有为该内部局域网主机建立地址转换映射项，则 NAT 服务器会创建一个会话，并给该会话分配一个端口。之后将源地址及端口改为内部全局 IP 地址及相应的端口，并发送数据包到外网主机或服务器上。接着，外网主机接收到信息后，将应答信息返回给端口 NAT 服务器。当端口 NAT 服务器接收到外网主机的应答信息后，就检查端口 NAT 映射表（图 8.1 中的 NAT 转换表），如果映射表存在匹配的映射项，则将目标地址及端口转换为对应的内网 IP 及端口，将数据包转发给内网主机；如不存在匹配项则丢弃。

后续大部分有关配置的例子都是围绕图 8.1 讲解的。

2. 静态 NAT 地址转换

【例 12】 通过内部本地地址（192.168.1.2/28～192.168.1.4/28）转换为内部全局地址（68.192.160.2/29～68.192.160.4/29），说明静态 NAT 地址转换。具体配置步骤如下。

答：（1）设置外部接口。

```
RouterA(config)# interface Serial 0                    !设置 s0 接口连接外部网络
RouterA(config-if)# ip address 68.192.160.1 255.255.255.248
RouterA(config-if)# ip nat outside
RouterA(config-if)# exit
```

（2）设置内部接口。

```
RouterA(config)# interface FastEthernet 0              !设置 f0 接口连接内部网络
RouterA(config-if)# ip address 192.168.1.1 255.255.255.240
RouterA(config-if)# ip nat inside
RouterA(config-if)# exit
```

（3）在内部本地与内部全局地址之间建立静态地址转换。

```
RouterA(config)# ip nat inside source static 192.168.1.2 68.192.160.2    !静态地址转换
```

```
RouterA(config)#ip nat inside source static 192.168.1.3 68.192.160.3    !静态地址转换
RouterA(config)#ip nat inside source static 192.168.1.4 68.192.160.4    !静态地址转换
RouterA(config-if)#exit
RouterA#show ip nat translation                            !显示地址转换有关信息
```

【例 13】　内部网络中有 WWW、E-mail 和 FTP 三台服务器，使用的内部全局地址分别为 68.192.160.2、68.192.160.3 和 68.192.160.4，路由器 f0 口的 IP 地址为 192.168.1.1，s0 口的 IP 地址为 68.192.160.1。在 Internet 上的主机都允许访问这三服务器，要求通过相应的静态地址转换配置，实现网络要求功能。

答：静态地址转换配置如下：

```
RouterA#configure terminal
RouterA(config)#ip nat inside source static 192.168.1.2 68.192.160.2 !WWW 静态地址转换
RouterA(config)#ip nat inside source static 192.168.1.3 68.192.160.3 !E-mail 静态地址转换
RouterA(config)#ip nat inside source static 192.168.1.4 68.192.160.4 !FTP 静态地址转换
RouterA(config)#interface FastEthernet 0                 !设置 f0 为内部转换接口
RouterA(config)#ip address 192.168.1.1 255.255.255.0
RouterA(config)#ip nat inside
RouterA(config)#no shutdown
RouterA(config)#interface Serial 0                       !设置 s0 为外部转换接口
RouterA(config)#ip address 68.192.160.1 255.255.255.0
RouterA(config)#ip nat outside
RouterA(config)#no shutdown
```

3. 动态 NAT 配置

动态 NAT 一般用于局域网中多个内部本地地址 IP 在进/出口时，从内部全局地址 IP 地址池中提取合法的公有 IP 地址对外联系。

【例 14】　如图 8.1 所示，内网主机 PC1(IP 为 192.168.2/28)，需要使用端口 3105 访问公网上的 Web 服务器 202.113.64.2(端口为 80)。

答：1) 访问过程

(1) 内网主机 PC1 将请求数据包传送给 NAT 服务器，数据包包头信息如下。

源 IP：192.168.1.2

源端口：3105

目标 IP：202.113.64.2

目标端口：80

(2) NAT 服务器收到数据包后，将数据源地址改为 NAT 服务器的内部全局 IP 地址，并将端口改为一个未使用的端口号 5234，这时数据包包头信息如下。

源 IP：68.192.160.2

源端口：5234

目标 IP：202.113.64.2

目标端口：80

同时，NAT 服务器建立一个 NAT 表，记录数据包包头信息的变更情况。

源 IP：192.168.1.2

源端口：3105

目标 IP：68.192.160.2

目标端口：5234

(3) Web 服务器收到数据包后，根据源 IP 和端口将返回内容传送给 NAT 服务器，数据包包头信息如下。

源 IP：202.113.64.2

源端口：80

目标 IP：68.192.160.2

目标端口：5234

(4) NAT 服务器接收到 Web 服务器发过来的数据包后，根据之前建立的 NAT 表，变更 IP 数据包信息。

源 IP：68.192.160.2

源端口：5234

目标 IP：192.168.1.2

目标端口：3105

2) 配置过程

通过以上步骤，明白了内私网主机与外部服务器间数据交互过程。以下将以 RouterA 作为网关设备，内部全局 IP 地址设置为一个地址池，名为 aaa，当内部网络要访问外网时，从 aaa 提取公网 IP。说明端口 NAT 转换配置。配置步骤如下。

(1) 设置外部端口。

```
RouterA(config)#interface Serial 0
RouterA(config-if)#ip address 68.192.160.1 255.255.255.248     !配置外部接口对应 IP
RouterA(config-if)#ip nat outside                              !s0 确定为外部网络接口
RouterA(config-if)#exit
```

(2) 设置内部接口。

```
RouterA(config)#int FastEthernet 0
RouterA(config-if)#ip address 192.168.1.1 255.255.255.240      !路由器的内部接口 IP
RouterA(config-if)#ip nat inside                               !f0 确定为内部网络接口
RouterA(config-if)#exit
```

(3) 定义内部全局 IP 地址池 aaa，将内部全局 IP 地址 68.192.160.2/29 放入 aaa。

```
RouterA (config)#ip nat pool aaa 68.192.160.2 68.192.160.6 netmask 255.255.255.248
```

(4) 定义内部访问列表 1，允许范围是内部 IP 地址 192.168.1.1～192.168.1.14/28。

```
RouterA (config)#access-list 1 permit 192.168.1.0 0.0.0.15
```

(5) 设置复用动态地址转换，实现内部本地与内部全局地址的转换。

```
RouterA(config)#ip nat inside source list 1 pool aaa
RouterA(config)#ip route 0.0.0.0 0.0.0.0 s0                    !配置默认路由
RouterA(config-if)#exit
RouterA#show ip nat translation
```

4. 端口地址转换配置

端口地址转换(Port Address Translate,PAT)是一种特殊动态 NAT,它用于将多个内部本地 IP 地址映射到一个公网 IP 的不同端口上。将原动态 nat 命令行地址池 pool 改变成为对外接口,并在后边加上参数 overload。

【例 15】 局域网使用 f0 接口的内部本地地址 192.168.1.0/28,要求通过地址端口转换配置,以实现局域网与 Internet 的通信。在内部网络中只有一个内部全局地址 68.192.160.1,在路由器的 s0 口。

答:地址端口转换配置如下。

```
RouterA#configure terminal
RouterA(config)#access-list 1 permit 192.168.1.0 0.0.0.15      !配置访问控制列表 1
RouterA(config)#ip nat inside source list 1 interface Serial 0 overload
                                                                !内部地址映射到 s0
RouterA(config)#ip route 0.0.0.0 0.0.0.0 Serial 0              !配置默认路由
RouterA(config)#inteface FastEthernet 0
RouterA(config-if)#ip address 192.168.1.1 255.255.255.240
RouterA(config-if)#ip nat inside                                !f0 确定为内部网络接口
RouterA(config)#no shutdown
RouterA(config-if)#exit
RouterA(config)#interface Serial 0
RouterA(config-if)#ip address 68.192.160.1 255.255.255.248
RouterA(config-if)#ip nat outside                               !s0 确定为外部网络接口
RouterA(config)#no shutdown
RouterA(config-if)#end
RouterA#show ip nat translation
```

5. 基于 NAT 的负载均衡

一般的 NAT 都是将内部私有 IP 转换为自己的公网 IP,负载均衡是将一个公网 IP 翻译成多个内部私有 IP。例如内部的 WWW 负载重时,可设置成多台服务器,各有自己的私有 IP,但对外还是映射成为一个统一的 IP 地址,对外部来讲多台服务器是捆绑在一起的一个虚拟服务器,外部访问这个虚拟服务器时,轮流指向各台服务器,从而达到负载均衡。

【例 16】 将来自内部局域网(192.168.1.0/28)的主机放入地址池(aaa)中,可以选择一个地址(如 68.192.160.2)作为自己的合法地址,经由 s0 口访问 Internet,实现负载均衡。

答:具体配置如下。

```
RouterA(config)#interface FastEthernet 0
RouterA(config-if)#ip address 192.168.1.1 255.255.255.240
RouterA(config-if)#ip nat inside
RouterA(config-if)#exit
RouterA(config)#interface Serial 0
RouterA(config-if)#ip address 68.192.160.1 255.255.255.248
RouterA(config-if)#ip nat outside
RouterA(config-if)#exit
RouterA(config)#ip nat pool aaa 192.168.1.2 192.168.1.14 netmask 255.255.255.240 type rotary
                         !type rotary 用来实现负载均衡,可能有些版本的模拟器不支持设置该参数
RouterA(config)#access-list 1 permit 68.192.160.2  !access-list 1 允许一个对外的 IP 地址
RouterA(config)#ip nat inside destination list 1 pool aaa     !内部地址映射成一个统一的 IP
```

```
RouterA(config)#ip route 0.0.0.0 0.0.0.0 Serial 0
RouterA(config-if)#exit
RouterA#show ip nat translation
```

本配置将 IP 地址 192.168.1.2/28～192.168.1.14/28 放入地址池 aaa，采用轮询方式，达到负载均衡，其中 rotary 参数是轮流的意思；list 1 指定的 IP 地址 68.192.160.2 就是被访问虚拟服务器的 IP 地址；inside destination 表示内部全局目的地址。

6. 基于服务的 NAT 配置

基于服务的 NAT 地址转换配置可以具体到协议和端口，以下举例说明。

【例 17】 假设内部地址 192.168.1.2/28 是一个 FTP 服务器，TCP 端口为 20/21；内部地址 192.168.1.3/28 是一个 WWW 服务器，TCP 端口为 80。外部地址为 68.192.160.1/29。

答：基于服务的 NAT 具体配置如下。

```
RouterA(config)#interface FastEthernet 0
RouterA(config-if)#ip address 192.168.1.1 255.255.255.240
RouterA(config-if)#ip nat inside                              !设置 f0 接口连接内部网络
RouterA(config)#no shutdown
RouterA(config-if)#exit
RouterA(config)#interface Serial 0
RouterA(config-if)#ip address 68.192.160.1 255.255.255.248
RouterA(config-if)#ip nat outside                             !设置 s0 接口连接外部网络
RouterA(config)#no shutdown
RouterA(config-if)#exit
RouterA(config)#ip nat inside source static tcp 192.168.1.2 20 68.192.160.1 20   !tcp 转换
RouterA(config)#ip nat inside source static tcp 192.168.1.2 21 68.192.160.1 21   !tcp 转换
RouterA(config)#ip nat inside source static tcp 192.168.1.3 80 68.192.160.1 80   !www 转换
RouterA(config)#ip route 0.0.0.0 0.0.0.0 s0
RouterA(config)#exit
RouterA#show ip nat translation
```

TCP(192.168.1.2)和 WWW(192.168.1.3)对外部来说都是一个 IP 地址为 68.192.160.1 的虚拟服务器，外部对 68.192.160.1 的 21 端口的访问会转到 192.168.1.2 服务器上，而对于 68.192.160.1 的 80 端口的访问将转到 192.168.1.3 这台服务器上。这里，由于目的地址端口被指定，所以 NAT 可以直接使用接口的 IP 地址。

7. 基于 NAT 的允许多个内部地址使用相同的内部全局地址

命令 ip nat inside source list 1 pool aaa overload 中的参数 overload 是允许多个内部主机地址使用相同的内部全局地址(如：68.192.160.2)。主要设置如下。

```
RouterA(config)#ip nat pool aaa 192.168.1.1 192.168.1.14 netmask 255.255.255.240
RouterA(config)#access-list 1 permit 68.192.160.2   !list 1 允许 68.192.160.2 流量通过
RouterA(config)#ip nat inside source list 1 pool aaa overload   !内部地址使用 68.192.160.2
```

8.3 分组交换(X.25)配置

8.3.1 X.25 基本配置

要在一个设备的端口上配置 X.25，需要做以下操作：对接口进行封装，设置参数，设置

接口的 X.121 地址和配置 X.25 MAP 等。以下给出的配置 X.25 命令采用华为设备。

1. X.25 基本配置命令

(1) 配置 X.25 工作模式。

link - protocol x25 {[dte | dce] | [nonstandardietf]}

支持 X.25 功能系列路由器所支持的 X.25 第三层可以工作在 DTE 模式,也可以工作在 DCE 模式,同时还可以指定进行数据报封装的格式,可选择的封装格式有 Cisco 兼容、DDN 和 IETF 等格式。如果只是简单地将两台路由器的一对串行接口背靠背直连进行数据传输,此时只要保证传输的两端分别为 DTE 和 DCE,并且封装格式一致即可。

(2) 配置接口的 X.121 地址。

x25 x121 - address x.121 - address

如:[Quidway-Serial0]x25 x121-address 20112451

X.121 规范中定义了 X.25 地址的格式,即 X.121 地址。当前地址的最大长度是 15 位数字。典型的地址长度是 12 位或 14 位数字。一般情况下,用户只需要为按 IETF 格式及 Cisco 兼容格式封装的 X.25 接口指定 X.121 地址。

(3) 创建远端协议地址到 X.121 地址的映射。

x25 map protocol protocol - address **x121 - address** x.121 - address [option]

如:[Quidway-Serial0]x25 map ip 10.0.0.1 x121-address 20112451

对于一个 X.25 接口,首先它拥有自己的 X.121 地址,并且拥有自己的网际协议地址(如 IP 协议)。当 X.25 通过这个接口发起呼叫时,它在呼叫请求分组中携带的源地址(即主叫 DTE 地址)就是这个接口的 X.121 地址。

protocol-address 和 x.121-address 指的是目的地的协议地址和 X.121 地址,而非本地的;对于每一个目的地,都需要创建一条地址映射。

(4) 创建永久虚电路。

x25 pvc pvc - number **protocol** protocol - address **x121 - address** x.121 - address

如:[Quidway-Serial0]x25 pvc 123 ip 10.0.0.1 x121-address 20112451

创建永久虚电路同时隐含地创建一条地址映射。对于数据流量大,通过租用专线连接的数据传输要求,可以为其创建永久虚电路。永久虚电路不需要经过呼叫过程,并且始终存在。而且在创建永久虚电路之前,不必先创建地址映射,因为在创建永久虚电路的同时,已经隐含地创建了一条地址映射。

2. X.25 的分组参数配置

1) 配置 X.25 虚电路范围

x25 vc - range {in - channel hic lic | bi - channel htc ltc | out - channel hoc loc}

X.25 协议可以将 DTE/DCE 之间的一条实际的物理链路复用,建立多条在逻辑上存在的虚连接,这种虚连接称为虚电路(Virtual Circuit,VC)或逻辑信道(Logic Channel,LC)。X.25 可以建立的虚连接最多可达 4095 条,编号从 1~4095。

X.25 协议中很重要的一部分内容就是如何管理这 4095 条虚电路。所有的虚电路号被划分成 4 个区域,X.25 协议使用 6 个参数来界定这 4 个区域,如图 8.2 所示。

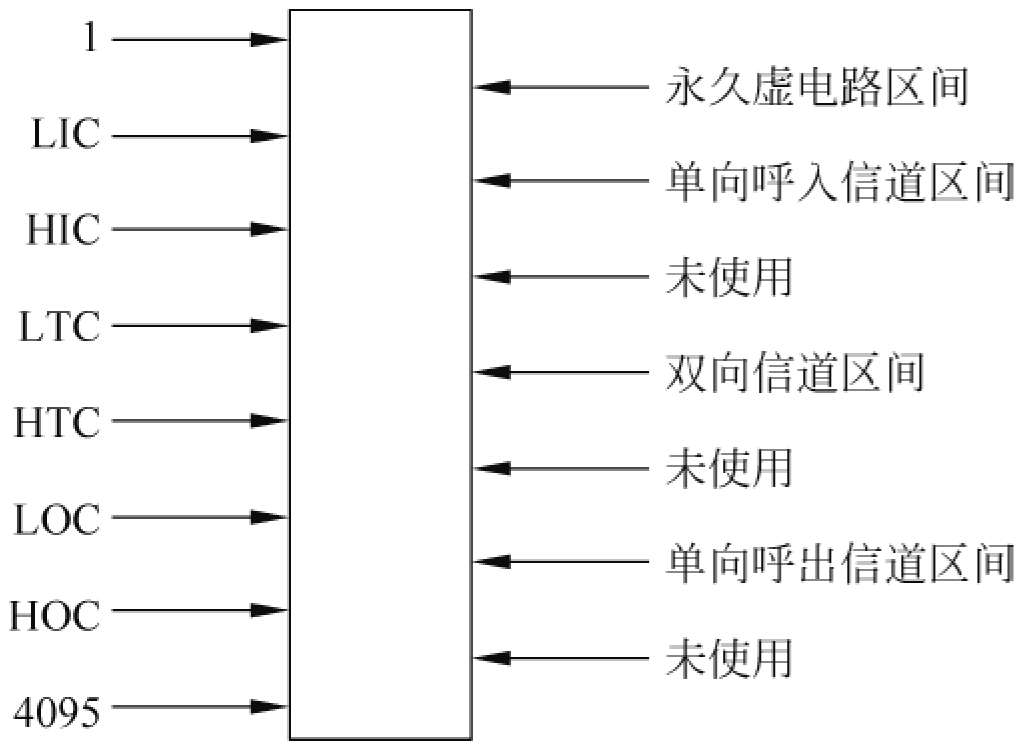


图 8.2 X.25 协议使用虚电路号被划分成 4 个区域

每一个区间(永久虚电路区间除外)被两个参数定义,可以称其为该区间的上限和下限;每个参数均可在 1～4095 之间(包括 1 和 4095)取值,但是,只有同时满足如下条件的配置才被认为是正确的配置:

严格升序,即 1≤LIC≤HIC<LTC≤HIC<LOC≤HOC≤4095;

若某个区间的上、下限其中之一为 0,那么另一个也必须为 0(上、下限均为 0 表示该区间被禁止使用)。

2) 配置默认的流量控制参数命令

```
x25 packet - size in - packets out - packets          !配置入、出分组大小,默认值为 128
x25 window - size in - packets out - packets          !配置入、出窗口大小,默认值为 2
```

X.25 协议是具有强流量控制能力的可靠传输协议,它具有这种能力的基础是“窗口尺寸”和“最大分组长度”。

3) 配置 X.25 分组编号模数

x25 modulo {8 | 128}:配置模编号方式,默认值为 8。

支持 X.25 功能系列路由器中 X.25 支持模 8 和模 128 两种分组顺序编号方式,模 8 方式是默认的编号方式。

因为 X.25 规程需要 DTE、DCE 两侧具有同样的分组顺序编号方式,所以完成配置后,需要执行 shutdown 与 undo shutdown 命令。配置操作过程如下。

(1) 进入主接口命令。

```
interface serial number
```

(2) 封装 X.25 协议命令。

```
link - protocol x25
```

(3) 创建虚拟子接口命令。

```
interface serial number subinterface - number [multipoint | point - to - point]
```

（4）配置地址映射命令。

x25 map protocol protocol－address **x121－address x.**121－address [option]

（5）配置永久虚电路命令。

x25 pvc pvc－number protocol protocol－address **x121－address** x.121－address [option]

X.25 子接口是一个虚拟接口，它有自己的协议地址和虚电路。在一个物理接口上可以创建多个子接口，这样就可以用一个物理接口实现多个网络的互连。X.25 的子接口又可以分为两种类型：点到点（point-to-point）子接口和点到多点（multipoint）子接口。点到点子接口用于连接单个远端，点到多点子接口用于连接多个远端，这些远端都必须在同一个网段。

8.3.2　X.25 典型配置举例

网络结构如图 8.3 所示，如果只是需要将两台路由器简单地背靠背连接，直连串口之间封装 X.25 协议并承载 IP 数据报进行传输，只要如下配置两台路由器即可。

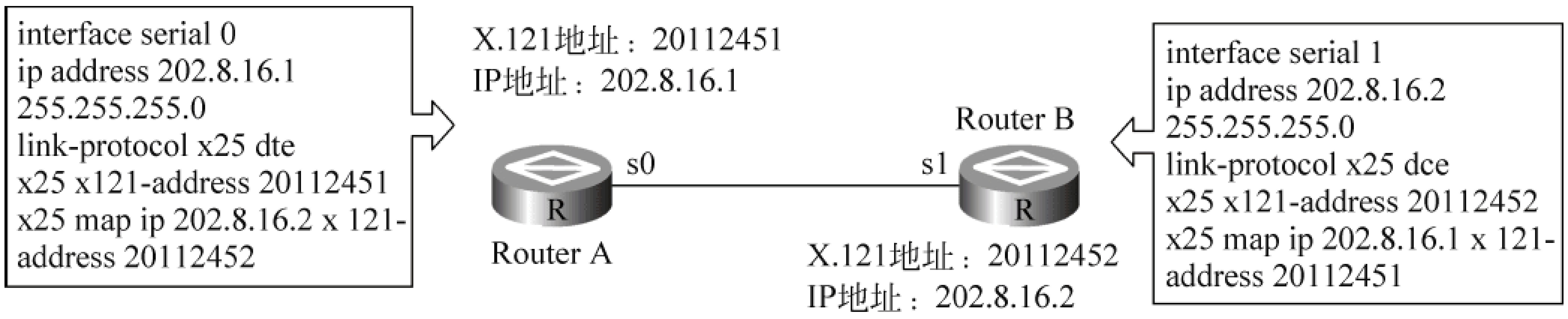


图 8.3　X.25 典型配置

1. 配置 RouterA

```
[Quidway]interface serial 0 !选定接口
[Quidway－Serial0]ip address 202.8.16.1 255.255.255.0 !为该接口指定 IP 地址
[Quidway－Serial0]link－protocol x25 dte          !将该接口封装为 X.25 接口，并指定为 DTE
[Quidway－Serial0]x25 x121－address 20112451      !指定该接口的 X.121 地址
[Quidway－Serial0]x25 map ip 202.8.16.2 x121－address 20112452
                                                  !到对端的地址映射
[Quidway－Serial0]x25 packet－size 1024 1024      !设置入、出流量控制参数
[Quidway－Serial0]x25 window－size 5 5            !设置入、出窗口参数
```

2. 配置 RouterB

```
[Quidway]interface serial 1 !选定接口
[Quidway－Serial1]ip add 202.8.16.2 255.255.255.0  !为该接口指定 IP 地址
[Quidway－Serial1]link－protocol x25 dce          !将该接口封装为 X.25 接口，并指定为 DCE
[Quidway－Serial1]x25 x121－address 20112452      !指定该接口的 X.121 地址
[Quidway－Serial0]x25 map ip 202.8.16.1 x121－address 20112451
                                                  !指定到对端的地址映射
[Quidway－Serial0]x25 packet－size 1024 1024      !设置入、出流量控制参数
[Quidway－Serial0]x25 window－size 5 5            !设置入、出窗口参数
```

8.4 帧中继配置

8.4.1 FR 配置命令

1. 配置 FR 的华为(Quidway)命令

(1) 封装帧中继协议。

命令：**link - protocol frame - relay** [nonstandard | ietf]

(2) 配置帧中继接口的终端类型。

命令：**frame - relay interface - type** [dce | dte | nni]

(3) 选择 LMI 类型。

命令：**frame - relay lmi - type** [ansi | cisco - compatible | q933a]

其中 lmi(Local Management Interface)为本地管理接口、ansi(American National Standards Institute)是美国国家标准学会指定的标准。

(4) 配置帧中继静态地址映射。

命令：**frame - relay map** {ip | ipx} protocol - address dlci [broadcast]

当对端路由器不支持逆向动态地址解析协议时，必须配置帧中继的静态地址映射。broadcast 的作用是在该接口上发送广播信息。

(5) 允许/禁止动态逆向地址解析。

命令：**frame - relay inarp** [ip | ipx] [dlci]

在运行了逆向地址解析协议(Inverse ARP)后，就能动态建立对端协议地址与本地 DLCI 的映射关系，适用于对端路由器也支持“逆向地址解析协议”或是网络环境较复杂的情况。

(6) 配置帧中继本地虚电路。

命令：**frame - relay dlci** dlci

为主接口和子接口分配一条虚电路号。虚电路号是本地有效的，也就是说，链路两端的虚电路号是可以相同的。也可为多个接口指定相同的虚电路号，但在一个物理接口上，虚电路号必须是唯一的。

(7) 配置帧中继子接口。

命令：**interface type** number. subinterface - number [multipoint | point - to - point]

number 指明了物理接口号；subinterface-number 是子接口号，如 s2/0.2，表示串行接口 2/0 的子接口为 2。子接口是一个逻辑结构，可以配置协议地址和虚电路 PVC 等，一个物理接口可以有多个子接口。虽然子接口是逻辑结构，并不实际存在，但对于网络层而言，子接口和主接口没有区别，都可通过配置 PVC 与远端设备相连。

帧中继子接口有点到点(point-to-point)和点到多点(multipoint)两类。点到点子接口用于连接单个对端，点到多点子接口用于连接同一个网段的多个对端。

2. 配置 FR 的 Cisco 命令

(1) 开启路由器的帧中继交换功能，用路由器来模拟 FR 交换机。

如：R(config)# frame-relay switching

(2) 设置时钟速率,只是 DCE 端才设置。

如:R(config-if)#clock rate 1000000

(3) 选择封装的 FR,通常使用国际标准 IETF 的标准封装,也是默认的。

如:R(config-if)#encapsulation frame-relay ietf

(4) 设置 FR 的本地管理接口类型,通常标准有 cisco、ansi、q933a,可以选择其中一种,但要注意其他与其通信的设备必须选同一种标准。

如:R(config-if)#frame-relay lmi-type q933a

(5) 设置接口类型,可选三种类型 dce、dte、nni。

如:R(config-if)#frame-relay intf-type dce

(6) 配置 FR 的一端到另一端的帧中继路由。

如:R(config-if)#frame-relay route 101 interface s0/2 201

如果本配置进入的是 s0/0 接口,则可以允许 DLCI 为 101 的数据从 s0/0 接口进入,然后从该路由器的 s0/2 接口出去后封装 DLCI 为 201。

(7) 配置 FR 的点对点子接口模式。

如:R(config)#interface s2/0.2 point-to-point

(8) 配置 FR 的接口封装 DLCI 号。

如:R(config-subif)#frame-relay interface-dlci 102

(9) 配置 FR 的本地接口 DLCI 号与对端的 IP 地址映射。

如:R(config-if)#frame-relay map ip 192.168.1.1 301 broadcast

8.4.2　FR 典型配置实例

1. 通过路由器设置帧中继专线

通过两台华为(Quidway)路由器的串口直连,Router A 工作在帧中继的 DCE 方式,Router B 工作在帧中继的 DTE 方式。实现帧中继专线互连局域网如图 8.4 所示,配置步骤如下。

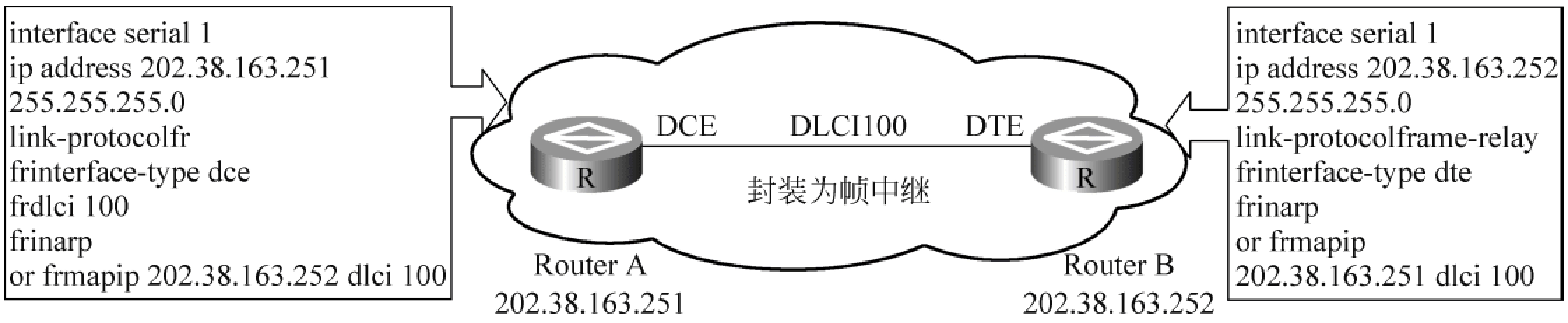


图 8.4　实现帧中继专线互连

(1) 配置 Router A。

```
[Quidway]fr switching                                  !在全局配置模式下配置帧中继交换
[Quidway]interface serial 1                            !进入 s1
[Quidway-Serial1]ip address 202.38.163.251 255.255.255.0    !配置 s1 接口 IP 地址
[Quidway-Serial1]link-protocol fr                      !S1 接口封装为帧中继
[Quidway-Serial1]fr interface-type dce                 !设置本端为 DCE
[Quidway-Serial1]fr dlci 100                           !配置本段虚电路,即 dlci 为 100
```

```
[Quidway-Serial1]fr inarp                    !如果对端路由器支持逆向地址解析,则配置动态地址映射
[Quidway-Serial1]fr map ip 202.38.163.252 dlci 100   !否则配置静态地址映射
```

(2) 配置 Router B。

```
[Quidway]interface serial 1                          !进入 s1 接口
[Quidway-Serial1]ip address 202.38.163.252 255.255.255.0
                                                     !设置 IP 地址
[Quidway-Serial1]link-protocol fr                    !该接口封装为帧中继
[Quidway-Serial1]fr interface-type dte               !设置本端为 DTE
[Quidway-Serial1]fr dlci 100                 !因接口类型为 DTE 时,又不是子接口,此命令可以不配
[Quidway-Serial1]fr inarp                    !如果对端路由器支持逆向地址解析,则配置动态地址映射
[Quidway-Serial1]fr map ip 202.38.163.251 dlci 100   !否则配置静态地址映射
```

2. 通过帧中继网络互连局域网

在通过公用帧中继网络互连局域网时,路由器只能作为帧中继的 DTE。图 8.5 给出了通过帧中继网络互连的局域网,表 8.1 给出了各个 Cisco 路由器对应设置要求。

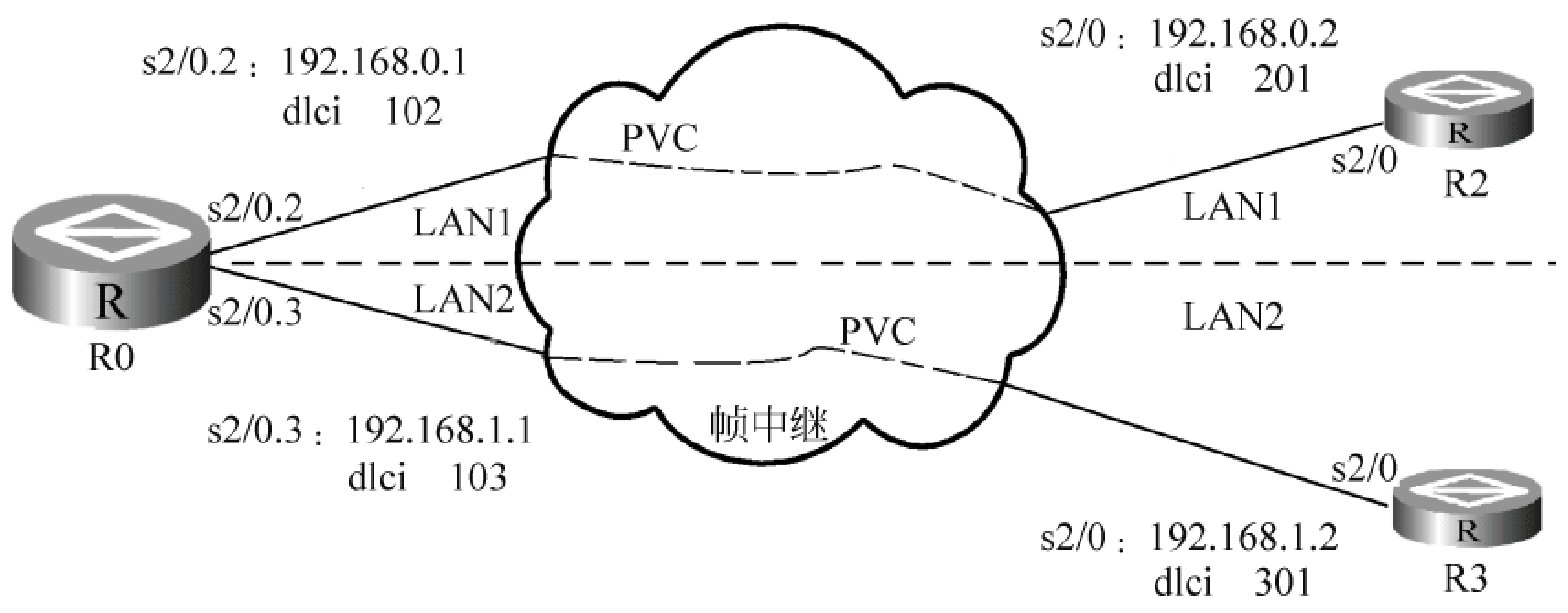


图 8.5 通过帧中继网络互连局域网

表 8.1 Cisco 路由器对应设置要求

路由器	路由器串行接口	接口对应 IP 地址	接口对应 dlci 号	接口所属网段
R0	s2/0.2(子接口 2)	192.168.0.1	102	LAN1
R0	s2/0.3(子接口 3)	192.168.1.1	103	LAN2
R2	s2/0	192.168.0.2	201	LAN1
R3	s2/0	192.168.1.2	301	LAN2

(1) R0 配置。

```
Router>enable
Router#config terminal
Router(config)#interface s2/0
Router(config-if)#encapsulation frame-relay          !接口 s2/0 封装为帧中继协议
Router(config-if)#no shutdown
Router(config-if)#exit
Router(config)#interface s2/0.2 point-to-point       !进入子接口 s2/0.2
Router(config-subif)#ip address 192.168.0.1 255.255.255.0      !s2/0.2 配置 IP,属于 LAN1
Router(config-subif)#frame-relay interface-dlci 102            !封装 DLCI 为 102
```

```
Router(config-if)#frame-relay map ip 192.168.0.2 102 broadcast   !在该接口上广播映射
Router(config-subif)#no shutdown
Router(config-subif)#exit
Router(config)#interface s2/0.3 point-to-point                   !进入子接口 s2/0.3
Router(config-subif)#ip address 192.168.1.1 255.255.255.0!s2/0.3
                                                                 !配置 IP,属于 LAN2
Router(config-subif)#frame-relay interface-dlci 103              !封装 DLCI 为 103
Router(config-if)#frame-relay map ip 192.168.1.2 103 broadcast   !本端 DLCI 与对端 IP 映射
Router(config-subif)#no shutdown
Router(config-subif)#exit
Router(config)#router rip                                        !进入 RIP 动态路由设置
Router(config-router)#network 192.168.0.0                        !设置到达 LAN1 的路由
Router(config-router)#network 192.168.1.0                        !设置到达 LAN2 的路由
```

（2）R2 配置。

```
Router>enable
Router#config terminal
Router(config)#interface s2/0
Router(config-if)#ip add 192.168.0.2 255.255.255.0               !s2/0.配置 IP,属于 LAN1
Router(config-if)#encapsulation frame-relay
Router(config-if)#frame-relay interface-dlci 201                 !封装 DLCI 为 201
Router(config-if)#frame-relay map ip 192.168.0.1 201 broadcast   !本端 DLCI 与对端 IP 映射
Router(config-subif)#no shutdown
Router(config-if)#exit
Router(config)#router rip                                        !进入 RIP 动态路由设置
Router(config-router)#network 192.168.1.0          !配置到达目标网 LAN2(192.168.1.0)路由
Router(config-router)#exit
Router(config-)#
```

（3）R3 配置。

```
Router>enable
Router#config terminal
Router(config)#interface s2/0
Router(config-if)#encapsulation frame-relay
Router(config-if)#ip add 192.168.1.2 255.255.255.0               !s2/0.配置 IP,属于 LAN2
Router(config-if)#frame-relay interface-dlci 301                 !封装 DLCI 为 301
Router(config-if)#frame-relay map ip 192.168.1.1 301 broadcast
                                                                 !本端 DLCI 与对端 IP 映射
Router(config-subif)#no shutdown
Router(config-if)#exit
Router(config)#router rip                                        !进入 RIP 动态路由设置
Router(config-router)#network 192.168.0.0          !配置到达目标网 LAN1(192.168.0.0)路由
Router(config-router)#exit
Router(config)#
```

习题

1. 简述静态 NAT 配置命令。

2. 某局域网使用内部本地地址为 192.168.1.0/24，要求通过对 RouterA 的动态地址转换配置，以实现局域网与 Internet 的通信。内部全局 IP 地址范围为 68.192.160.2～68.192.160.4。外部接口为 s0，内部接口为 f0。要求完成动态地址转换配置（网络结构参考图 8.1）。

3. 参考图 8.5，自行设置要求，通过路由器完成帧中继配置。

第9章 自治区域内部协议

内部网关协议(Interior Gateway Protocol,IGP)为自治系统之内的路由协议总称。在同一个自治系统内部,除了静态路由外,常用的动态路由协议有:基于距离矢量的路由信息协议(RIP),基于链路状态的中间系统到中间系统协议(IS-IS),以及开放式最短路径优先协议(OSPF)。RIP 通过 UDP 交换路由信息,被广泛应用于区域性中、小型网络中,而与 OSPF 恰好形成互补;OSPF 是一种典型的链路状态路由协议,具有快速收敛性等特点;IS-IS 最早是 OSI 的无连接网络协议,后来通过修改应用到了 TCP/IP 网络环境中,它采用骨干区域与非骨干区域的两级分层结构来支持大规模的路由网络。本章将重点介绍 RIP、OSPF 的运行原理以及路由配置,并对 IS-IS 也作适当介绍。

9.1 路由信息协议

路由信息协议(Routing Information Protocol,RIP)共有 RIP-1、RIP-2 两个版本。RIP-2 在 RIP-1 协议的基础上增加了一些扩展特性,增加了支持变长子网掩码(VLSM),同时支持明文认证和 MD5 密文认证,有广播和组播两种方式,默认时将采用组播发送报文,RIP-2 的组播地址为 224.0.0.9。组播发送报文的优点是在同一网络中那些未运行 RIP 的主机可以避免接收 RIP 的广播报文。

9.1.1 RIP 报文结构

如图 9.1 所示为 RIPv2 报文结构,RIPv2 比 RIPv1 多了 Route Tag、SubNet Mask 和 Next Hop,以适用于现代网络路由选择环境。RIPv2 规范允许 RIP 报文包含更多的信息。报文结构内容如下。

8位	8位	16位
Command	Version	Unused(Set to zero)
Address Family Identifier		Route Tag
IP Address		
Subnet Mask		
Next Hop		
Metric		

图 9.1 RIPv2 报文结构

(1) 命令(Command):表示该报文是请求还是响应。为 1 时,表示 RIP 请求;为 2 时,表示应答。请求报文要求路由器发送其路由表的全部或部分。响应报文可以是主动提供的周期性路由更新或对请求的响应。

(2) 版本(Version):版本号。如此值为2,表示RIPv2。

(3) 未使用(Unused):此值为0。

(4) 地址族标志(Address Family Identifier):指明使用的地址族,RIP设计用于携带多种不同协议的路由信息。每个项都有地址族标志来表明使用的地址类型。

(5) 路由标记(Route Tag):用来支持外部网关协议,它传递自治系统的标号给外部网关协议及边界网关协议。没有路由标记的路由器必须将0作为自己的路由标记对外广播。

(6) IP地址(IP Address):指明目标网的地址。

(7) 子网掩码(Subnet Mask):包含该项的子网掩码。如果此字段为0,则该项不指定子网掩码。

(8) 下一跳(Next Hop):指明下一跳的IP地址。

(9) 跳数(Metric):在RIP中,路由器到与它直接相连网络的跳数为0,通过一个路由器可以直达网络的跳数为1,其余以此类推。为限制收敛时间,RIP规定Metric取值为0~15的整数,大于或等于16的跳数被定义为无穷大,即目的网络或主机不可达。

9.1.2 RIP工作原理

实际上,RIP作为一个系统长驻进程存在于路由器中,它负责从网络中的其他路由器接收路由信息,对路由表作动态的维护,保证了路由器发送报文时选择正确的路由,同时也广播本路由器的路由信息到相邻的路由器。

RIP处于UDP的上层,RIP所接收的路由信息都封装在UDP的数据报中,RIP在520号端口上接收来自远程路由器的路由修改信息,并对本地的路由表做相应的修改,完成路由表的维护和更新,同时通知其他路由器。通过这种方式,可达到全局路由的同步。

更新定时器(30s):用于设置定期路由更新的时间间隔。

无效定时器(180s):路由器在认定一个路由成为无效路由之前所需要等待的时间。

保持定时器(180s):用于设置路由信息被抑制的时间。

刷新定时器(240s):用于设置某个路由成为无效路由并将它从路由表中删除的时间。

RIP每隔30s向外发送一次更新报文。如果路由器经过180s没有收到来自对端的路由更新报文,则将所有来自此路由器的路由信息标志为不可达,以该度量值在响应报文(response)中发布4次(120s),之后从路由表中清除;对本路由表中不存在的路由项,在度量值小于不可达(16)时,在路由表中增加该路由项;路由表中的每一路由项都对应一个老化定时器,当路由项在180s内没有任何更新时,定时器超时,该路由项的度量值变为不可达(16)。

如果10个路由器链形相连,在第一个路由器上有路由更新时(例如增加网络接口),最坏条件下可能经过9个30s才能将路由信息传输到第10个路由器。因为路由收敛过程中路由域的不稳定性可能导致大量数据沿低效路径传输,所以这样长的收敛时间是无法接受的。

RIP路由自环问题如图9.2所示。当196.168.6.0网络不可达(例如路由器RouterA的端口停止工作)后,路由器RouterA将路由表中196.168.6.0的距离标为16,表示不可达。如果在路由器RouterA发送路由表更新以前,路由器RouterB将路由表发送到RouterA。则路由器RouterA会认为从路由器RouterB经过距离3可以到达网络196.168.6.0。下一次路由更新中,路由器RouterB会认为从路由器RouterA经过距离4可以到达网络

196.168.6.0。于是形成路由循环。在路由距离到达表示不可达的 16 以前,所有发往 196.168.6.0的数据包将在路由器 RouterA 和路由器 RouterB 之间循环打转。

图 9.2　RIP 路由自环问题

为防止产生路由环路,RIP 支持水平分割(split horizon)与路由中毒(poison reverse),并在路由中毒时采用触发更新(triggered update)。另外,RIP 还允许引入其他路由协议所得到的路由。

水平分割:路由器向某一路由条目的下一跳路由器发送本条路由是没有意义的。如果路由器 RouterB 的路由表中 196.168.6.0 的下一跳是路由器 RouterA,那么路由器 RouterB 没有必要向路由器 RouterA 发送 196.168.6.0 的路由信息。因为路由器 RouterB 的路由 196.168.6.0 是从路由器 RouterA 学到的,路由器 RouterA 无疑已经知道 196.168.6.0 这一条路由。因此不会出现路由循环问题,但是水平分割并不能解决所有循环问题。

水平分割毒化反转(路由中毒):当某路由不可达以后,路由器向相邻路由器广播该路由的不可达消息:距离为 16。当 196.168.6.0 网络不可达后,路由器 RouterA 通知路由器 RouterB,路由器 RouterB 又通知路由器 RouterC:到目标 196.168.6.0 网络距离为 16。还要进行反向再通知,路由器 RouterC 通知路由器 RouterB,路由器 RouterB 又通知路由器 RouterA:到目标 196.168.6.0 网络距离为 16。

即时更新:为解决上述问题,一个简单的措施就是在 30s 定时更新之外再加上路由立即更新。除每 30s 定期发送路由表外,一旦路由条目发生变化则立即发送更新信息。

【例 1】 如图 9.3 所示,路由器 RouterA 和 RouterB 直接相连,构成 3 个网段,即:196.168.64.0/24、211.138.137.0/24 和 172.16.0.0/24,要求说明网络运行路由协议 RIP 后的路由更新情况。

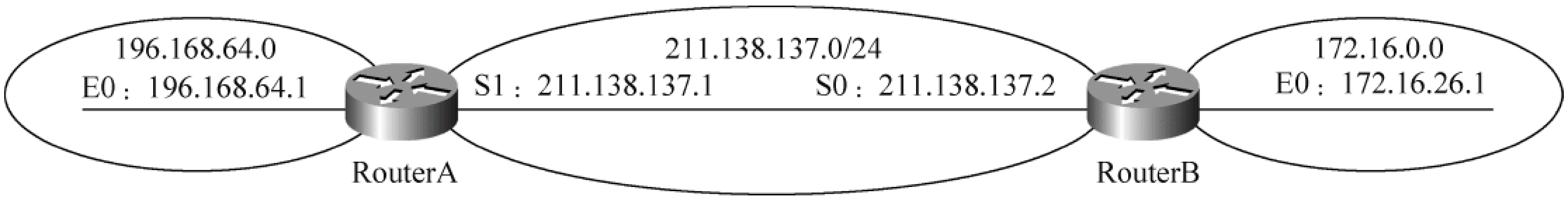


图 9.3　运行 RIP 的路由器与网段示意

答:在 RIP 中,每个路由器都在周期性地向其直连的邻居路由器发送自己完全的路由表,并且从自己直连的邻居路由器接收路由更新信息。首先看路由器 RouterA,对于网络 196.168.64.0/24,由于它和路由更新发出的接口 S1(211.138.137.1/24)不属于同一个主网络,所以,RIP 将子网 196.168.64.0 自动汇总为主网络:196.168.64.0/24,并且发送汇总后的主网络。对于网络 211.138.137.0/24,该网络路由更新将被发送。表 9.1 为路由器 RouterA 的路由表。

再看路由器 RouterB,首先,对于收到关于网络 211.138.137.0/24 的路由更新,由于它和接收路由更新的接口 S0(211.138.137.2/24)都属于同一个主网络,同时,路由表中已经有 211.138.137.0/24 网络的路由信息,而且是 S0 接口的直连路由信息,所以网络

211.138.137.0/24 的更新将被忽略。对于收到相邻路由器发来的网络 196.168.64.0/24，此路由将被安装。表 9.2 为路由器 RouterB 的路由表。

表 9.1 路由器 RouterA 的路由表

目的网络	输出接口	跳数
196.168.64.0	E0	0
211.138.137.0	S1	0
172.16.0.0	S1	1

表 9.2 路由器 RouterB 的路由表

目的网络	输出接口	跳数
211.138.137.0	S0	0
172.16.0.0	E0	0
196.168.64.0	S0	1

在每个 RIP 路由更新报文中，最多可以携带 25 个子网的路由信息。如果数量多于此值，则通过发送多个 RIP 报文来实现。

运行 RIP 的路由器在决定发送一条路由更新信息之前，首先要检查待发送路由更新的网络或子网是否和路由更新送出接口属于同一个网络。如果不属于同一个主网络，则 RIP 会将待发送路由更新的网络在主网络边界进行自动汇总。如果属于同一个主网络，而且两者的子网掩码相同则发送此路由更新，否则不发送此网络的路由更新信息。

当运行 RIP 的路由器收到一条路由更新消息后，首先要检查收到路由更新中的网络与接收更新接口是否属于同一个主网络。如果是，则路由更新中的网络子网掩码将使用接收更新的接口子网掩码。如果不属于同一主网络，同时路由更新中的网络子网已经存在于接收路由更新的路由表中了，而且是从另一接口收到的更新中学到的，则此路由更新被忽略。否则，安装此路由条目。

9.1.3 RIP 工作流程

RIP 路由器周期性地以多播形式向邻居发送自己的路由表复制件，即<目的，度量>组，每个接收到该消息的路由器修改消息中路由的度量，并在每条路由的度量上加上接收该路由消息接口的花费。通过图 9.4 可以看出 RIP 拥有如下简单的工作流程。

RIP 启动时的初始路由表仅包含本路由器的一些直连接口路由。RIP 启动后向各接口广播一个 Request 报文。

邻居路由器的 RIP 从某接口收到 Request 报文后，根据自己的路由表，形成 Response 报文向该接口对应的网络广播。

RIP 接收邻居路由器回复的包含邻居路由器路由表的 Response 报文，形成自己的路由表。RIP 根据 D-V 算法的特点，将协议的参加者分为主动机和被动机两种。主动机主动向外广播路由刷新报文，被动机被动地接收路由刷新报文。一般情况下，主机作为被动机，路由器则既是主动机又是被动机，即在向外广播路由刷新报文的同时，接收来自其他主动机的 D-V 报文，并进行路由刷新。RIP 以 30s 为周期用 Response 报文广播自己的路由表。

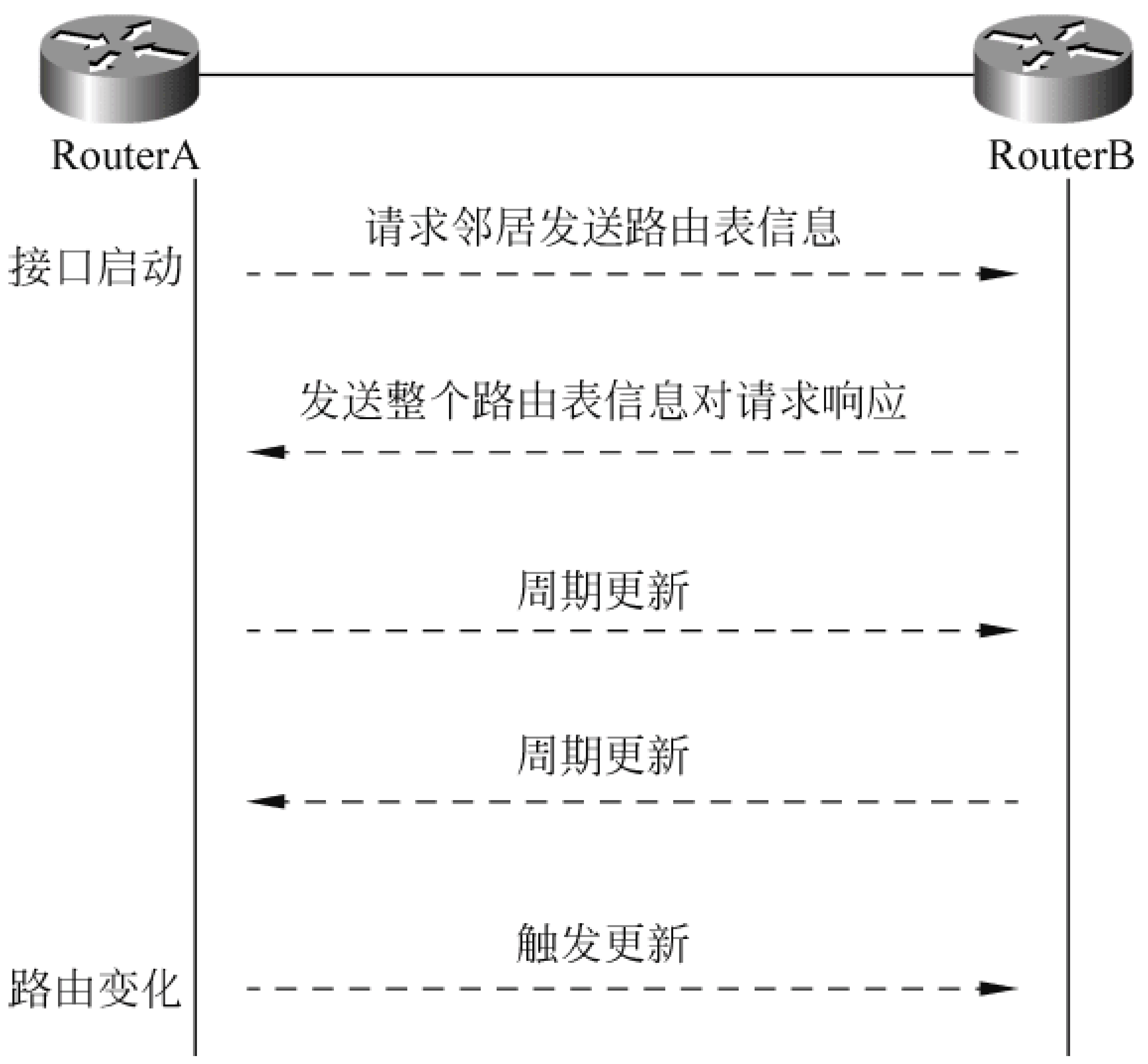


图 9.4　RIP 的工作流程

收到邻居发送来的 Response 报文后，RIP 计算报文中的路由项的度量值，比较其与本地路由表路由项度量值的差别，更新自己的路由表。

报文中路由项度量值的计算：metric'＝MIN(metric＋cost，16)，metric 为报文中携带的度量值信息，cost 为接收报文的网络度量值开销，默认为 1(表示 1 跳)，16 代表不可达。

9.1.4　RIP 基本配置

配置与接口相关的功能特性不受 RIP 是否使能的限制。需要注意的是，在关闭 RIP 后，原来的接口参数也同时失效。以下介绍的是华为路由器的命令及配置。

1. 启动 RIP

命令：**router rip**

要注意的是，有些路由器可以在全局配置模式下直接用 rip 命令启动 RIP，并进入 RIP 配置模式。要删除 rip，则要执行 no router rip。

2. 在指定网络上使能 RIP

RIP 任务启动后，还需要指定其工作网段，RIP 只能在指定网段上的接口工作。对于不在指定网段上的接口，RIP 既不在它上面接收和发送路由，也不将它的接口路由转发出去。

命令：

```
network{network-number | all}
```

当对某一地址使用命令 network 时，其效果是使能该地址的网段接口。network-number 为使能或不使能的网段地址，可为各个接口的 IP 网络地址。

3. 指定端口版本(接口模式下)

可指定接口所处理 RIP 报文的版本，命令如下。

```
rip version 1                              !启用 RIP 版本 1
rip version 2[bcast|mcast]                 !广播方式或多播方式
ip rip send | receive version 1 | 2        !更新路由器接口发送、接收 RIP 版本
```

其中,bcast(或 broadcast)为广播方式,mcast(或 multicast)为组播(或多播)方式。

4. 指定端口的工作状态(接口模式下)

可指定 RIP 在接口上的工作状态,如接口上是否运行 RIP,即是否在接口发送和接收 RIP 刷新报文,还可单独指定接口是否发送或者接收更新报文。命令如下:

```
rip work
rip input
rip output
```

rip work 从功能上等价于 rip input 与 rip output 两个命令。在默认情况下,一个接口可接收、发送 RIP 更新报文。

5. 配置 rip-2 路由聚合

命令: **auto - summary**

路由聚合(或称汇总)是指:同一自然网段内的不同子网的路由在向外(其他网段)发送时,可聚合成一条自然掩码的路由发送。路由聚合减少了路由表中的路由信息量。

6. 配置 rip-2 报文的认证(接口模式下)

命令:

```
rip authentication simple password              !配置 rip - 2 明文认证
rip authentication md5 key - string string      !配置 rip - 2 MD5 密文认证
```

只有 RIP-2 才进行报文的认证,RIP-2 支持两种认证方式:明文认证 Simple 和 MD5 密文认证。MD5 密文认证的报文格式有两种:一种遵循 RFC1723(RIP-2 Carrying Additional Information,携带附加信息)规定;另一种遵循 RFC2083(RIP-2 MD5 Authentication)规定。

【例 2】 如图 9.5 所示,RTA 和 RTB 之间链路层封装 PPP,RTB 和 RTC 之间链路层封装 FR 协议,所有路由器需要启动 RIP 路由协议,RTA 和 RTB 之间做 MD5 验证,要求对各路由器进行配置。

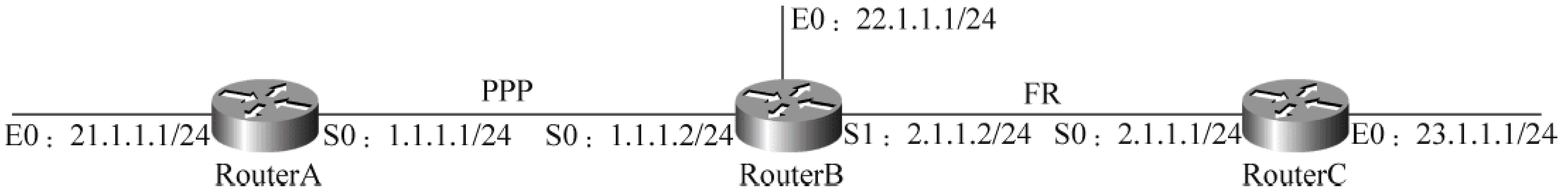


图 9.5 RTA 和 RTB 以及 RTC 之间的连接

答:(1) RouterA。

```
[RouterA]rip                                          !启动 rip
[RouterA - rip]network all                            !表示对 RouterA 所有接口使能
[RouterA - Ethernet0]rip version 2 broadcast          !指定接口 E0 版本号
[RouterA - Serial0]rip version 2 broadcast            !指定接口 S0 版本号
[RouterA - Serial0]rip authentication mode md5 key - string mu      !MD5 认证,mu 为密码
```

(2) RouterB。

```
[RouterB]rip                                          !启动 rip
[RouterB - rip]network all                            !表示对 RouterB 所有接口使能
[RouterB - rip]set peer 2.1.1.1                       !配置指定 VPN 链路对端接口的 IP 地址
[RouterB - Serial0]rip version2 broadcast             !指定接口 S0 版本号
```

```
[RouterB-Serial0]rip authentication-mode md5 key-string mu     !MD5 认证,mu 为密码
[RouterB-Serial1]rip version 2 broadcast     !指定接口 S1 版本号
[RouterB-Serial1]link-protocol fr     !配置接口封装为帧中继
```

(3) RouterC。

```
[RouterC]fr switching     !在全局配置模式下配置帧中继交换
[RouterC]rip     !启动 rip
[RouterC-rip]network all     !表示对 RouterC 所有接口使能
[RouterC-rip]set peer 2.1.1.2     !配置指定 VPN 链路对端接口的 IP 地址
[RouterC-Ethernet0]rip version2 broadcast     !指定接口 E0 版本号
[RouterC-Serial0]rip version 2 broadcast     !指定接口 S0 版本号
[RouterC-Serial0]link-protocol fr     !配置接口封装为帧中继
[RouterC-Serial0]fr interface-type DCE     !封装帧中继接口类型,设置为 DCE
[RouterC-Serial0]fr dlci 20     !配置本地虚电路,分配 DLCI(20)
```

(4) 用 display rip 显示当前 RIP 的运行状态如下。

```
RIP is turning on     !RIP 为运行状态
checkzero is on default-metric 16     !校验和开关打开,默认路由权为 16
no peer     !没有指定定点传送地址
network 1.1.0.0 2.1.0.0     !在 1.1.0.0 与 2.1.0.0 网段上使用 RIP 协议
summary is on preference 100     !自动聚合路由,RIP 路由的 preference 为 100
```

9.1.5 RIP 配置实例

以下给出两条 Cisco 路由器认证命令,还有些命令将在具体配置中加以说明。

```
ip rip authentication key-chain chain     !配置 rip-2 MD5 密文认证
rip authentication md5 type[nonstandard-compatible | usual]     !指定 MD5 类型
```

其中,nonstandard-compatible | usual 为非标准兼容|通常。

【例 3】 对于如图 9.6 所示网络,要求分别给出在路由器 RouterA、RouterB 和 RouterC 上配置 RIP 的命令过程。首先,启动 RIP 路由进程。然后,声明 RIP“关心”的网络。所谓“关心”的网络是指 RIP 以后将在属于这些网络的路由器接口上发送/接收路由更新信息。

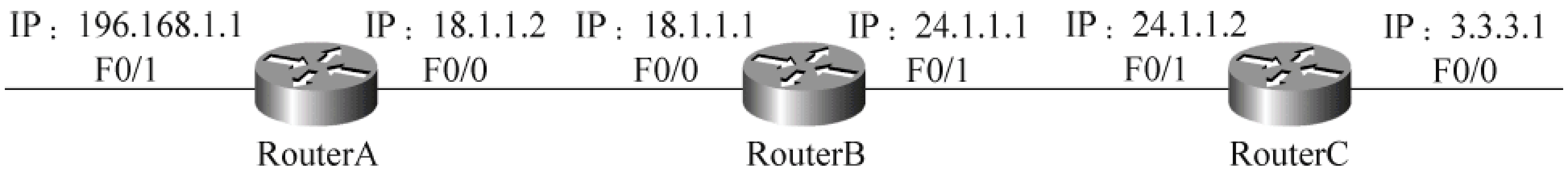


图 9.6 RIPv1 配置环境

答: (1) 在路由器 A 上配置 RIP 命令的过程。

```
RouterA#configure terminal
RouterA(config)#router rip
RouterA(config-router)#version 2
RouterA(config-router)#network 18.1.0.0     !指定 RouterA 的 f0/0 对应网络
RouterA(config-router)#network 196.168.0.0     !指定 RouterA 的 f0/1 对应网络
RouterA(config-router)#exit
```

(2) 在路由器 B 上配置 RIP 命令的过程。

```
RouterB#configure terminal
```

```
RouterB(config)#router rip
RouterB(config-router)#version 2
RouterB(config-router)#network 24.1.0.0          !指定 RouterB 的 f0/1 对应网络
RouterB(config-router)#network 18.1.0.0          !指定 RouterB 的 f0/0 对应网络
RouterB(config-router)#end
```

(3) 在路由器 C 上配置 RIP2 命令的过程。

```
RouterC#configure terminal
RouterC(config)#router rip
RouterC(config-router)#version 2
RouterC(config-router)#network 3.3.0.0           !指定 RouterC 的 f0/0 对应网络
RouterC(config-router)#network 24.1.0.0          !指定 RouterC 的 f0/1 对应网络
RouterC(config-router)#end
```

【例 4】 在图 9.7 中,RouterA、RouterB 之间的链路上启用 RIP 密文认证,要求实现完全配置,并显示相关信息。

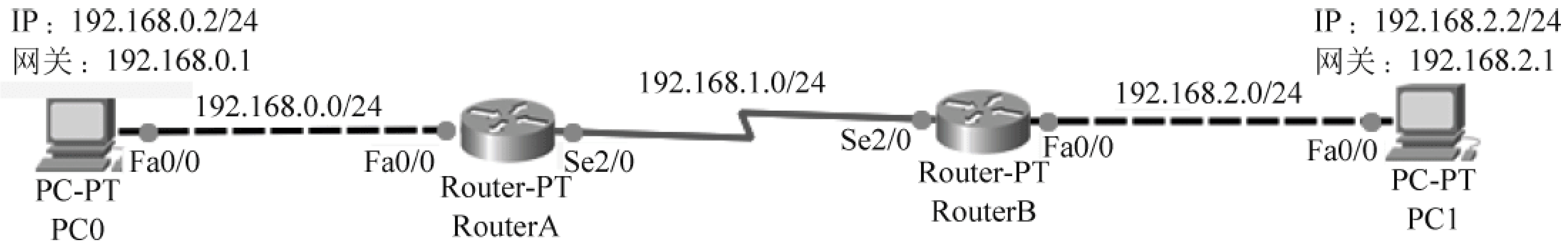


图 9.7 RIPv2 认证环境示意

答:(1) RouterA 的配置。

第一步:根据网络拓扑图的说明,配置各个接口的地址。

```
Router>
Router>enable
Router#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Router(config)#hostname RouterA
RouterA(config)#interface f0/0
RouterA(config-if)#ip address 192.168.0.1 255.255.255.0
RouterA(config-if)#no shutdown
%LINK-5-CHANGED: Interface FastEthernet0/0, changed state to up
RouterA(config-if)#exit
RouterA(config)#interface s2/0
RouterA(config-if)#ip address 192.168.1.1 255.255.255.0
RouterA(config-if)#no shutdown
%LINK-5-CHANGED: Interface Serial2/0, changed state to down
RouterA(config-if)#exit
```

第二步:上述配置完成后,接下来启动 RIPv2 的配置。以下给出 RouterA 的配置。

```
RouterA(config)#router rip
RouterA(config-router)#version 2
RouterA(config-router)#network 192.168.1.0
RouterA(config-router)#network 192.168.0.0
RouterA(config-router)#exit
```

第三步：上述配置完成后，接下来启动 RIPv2 认证的配置。下面给出在 RouterA 上对 RouterB 配置 md5 认证的步骤。

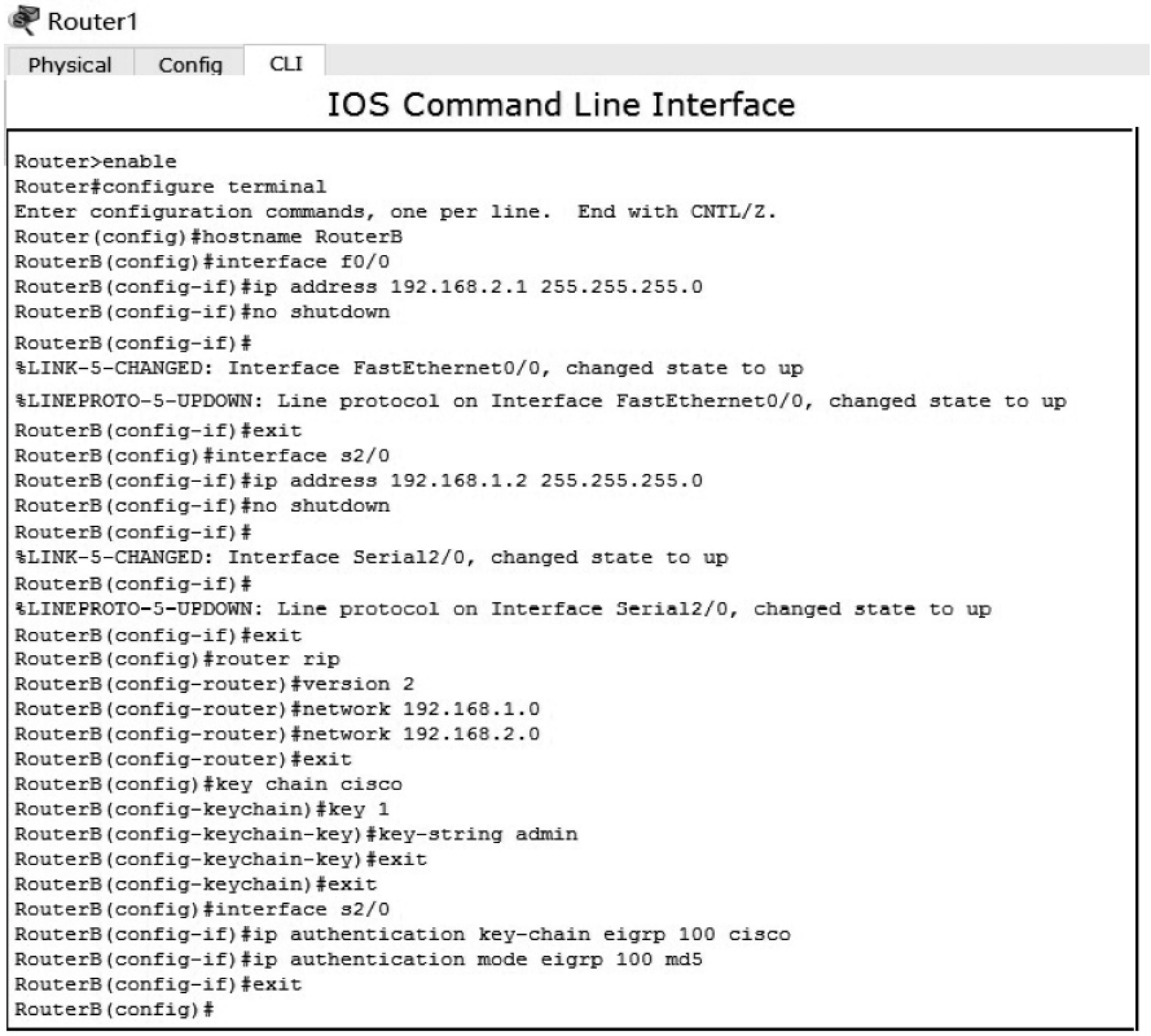
```
RouterA(config)#key chain cisco                              !定义一个密钥链(key chain)名字 cisco
RouterA(config-keychain)#key 1                               !定义一个(或一组)密钥链的钥匙(key)
RouterA(config-keychain-key)#key-string admin                !设置第一个密钥的值为 admin
RouterA(config-keychain-key)#exit                            !密钥值 admin 必须和路由器 B 的密钥值相同
RouterA(config-keychain)#exit
RouterA(config)#interface s2/0                               !在接口模式启用认证,并定义使用密钥链
RouterA(config-if)#ip rip authentication key-chain cisco     !密钥链为 cisco
RouterA(config-if)#ip authentication mode md5                !在接口模式启用 MD5 加密认证
```

需要说明的是以上两条命令在有些模拟器上可能不支持，但支持增强的内部网关路由选择协议(Enhanced Interior Gateway Routing Protocol，EIGRP)，所以也可用以下命令完成。

```
RouterA(config-if)#ip authentication key-chain eigrp 100 cisco
RouterA(config-if)#ip authentication mode eigrp 100 md5
RouterA(config-if)#exit
RouterA(config)#
```

(2) RouterB 的配置。

RouterB 的配置与 RouterA 类似，要注意网络地址的不同，而密钥值 admin 必须和路由器 A 的密钥值相同，具体配置如图 9.8 所示。


Router1
Physical　Config　CLI
IOS Command Line Interface

```
Router>enable
Router#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
Router(config)#hostname RouterB
RouterB(config)#interface f0/0
RouterB(config-if)#ip address 192.168.2.1 255.255.255.0
RouterB(config-if)#no shutdown
RouterB(config-if)#
%LINK-5-CHANGED: Interface FastEthernet0/0, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/0, changed state to up
RouterB(config-if)#exit
RouterB(config)#interface s2/0
RouterB(config-if)#ip address 192.168.1.2 255.255.255.0
RouterB(config-if)#no shutdown
RouterB(config-if)#
%LINK-5-CHANGED: Interface Serial2/0, changed state to up
RouterB(config-if)#
%LINEPROTO-5-UPDOWN: Line protocol on Interface Serial2/0, changed state to up
RouterB(config-if)#exit
RouterB(config)#router rip
RouterB(config-router)#version 2
RouterB(config-router)#network 192.168.1.0
RouterB(config-router)#network 192.168.2.0
RouterB(config-router)#exit
RouterB(config)#key chain cisco
RouterB(config-keychain)#key 1
RouterB(config-keychain-key)#key-string admin
RouterB(config-keychain-key)#exit
RouterB(config-keychain)#exit
RouterB(config)#interface s2/0
RouterB(config-if)#ip authentication key-chain eigrp 100 cisco
RouterB(config-if)#ip authentication mode eigrp 100 md5
RouterB(config-if)#exit
RouterB(config)#
```

图 9.8　RouterB 配置截图

在 RouterB 对 RouterA 的链路上也做相同的 RIP 认证配置后，RouterA 和 RouterB 之间就可以顺利交换路由更新了。

（3）设置 PC 并测试。

根据拓扑图，分别设置 PC0 的 IP 地址为 192.168.0.2/24，网关地址为 192.168.0.1；PC1 的 IP 地址为 192.168.2.2/24，网关地址为 192.168.2.1。然后再进行 PING 测试，其结果如图 9.9 所示，说明 RIP 在正常运作。

Fire	Last Status	Source	Destination	Type	Color	Time(sec)	Periodic	Num	Edit
●	Successful	PC1	Router0	ICMP		0.000	N	0	(edit)
●	Successful	PC1	PC0	ICMP		0.000	N	1	(edit)
●	Successful	PC0	PC1	ICMP		0.000	N	2	(edit)
●	Successful	Router0	PC1	ICMP		0.000	N	3	(edit)
●	Successful	PC0	192.168.2.1	ICMP		1.000	N	4	(edit)

图 9.9 PING 测试结果

（4）查看路由状况。

在路由器 RouterA 上执行“RouterA # show ip protocols”后，显示的 RIP 协议信息如图 9.10 所示，主要包含以下内容：

```
RouterA#show ip protocols
Routing Protocol is "rip"
Sending updates every 30 seconds, next due in 14 seconds
Invalid after 180 seconds, hold down 180, flushed after 240
Outgoing update filter list for all interfaces is not set
Incoming update filter list for all interfaces is not set
Redistributing: rip
Default version control: send version 2, receive 2
  Interface             Send  Recv  Triggered RIP  Key-chain
  Serial2/0             2     2
  FastEthernet0/0       2     2
Automatic network summarization is in effect
Maximum path: 4
Routing for Networks:
          192.168.0.0
          192.168.1.0
Passive Interface(s):
Routing Information Sources:
          Gateway         Distance      Last Update
          192.168.1.2          120      00:00:04
Distance: (default is 120)
RouterA#
```

图 9.10 RouterA 显示 RIP 协议信息

Routing Protocol is "rip"(路由协议是 RIP)

Sending updates every 30 seconds, next due in 14 seconds(每隔 30s 发送一次更新，且距离下一次更新还有 4s)

Invalid after 180 seconds, hold down 180, flushed after 240(180s 后无效，保持 down 为 180s，240s 后刷新)

Outgoing update filter list for all interfaces is not set(在出方向上对所有接口都没有设置过滤列表)

Incoming update filter list for all interfaces is not set(在入方向上对所有接口都没有设置过滤列表)

Redistributing: rip(重新分配：RIP)

Default version control: send version 2, receive 2(默认版本控制：发送版本 2，接收 2)

Interface Send Recv Triggered RIP Key－chain(接口发送/接收触发 RIP 的钥匙链)

Automatic network summarization is in effect(网络自动汇总生效)

Maximum path: 4(最大路径：4)

```
Routing for Networks:(网络路由)
Passive Interface(s):(被动接口):
Routing Information Sources: 路由信息源:
Gateway Distance Last Update(网关距离上次更新)
Distance: (default is 120)(距离: (默认值是 120))
```

执行 RouterA＃show ip route 后显示的部分路由信息如图 9.11 所示，其中："Gateway of last resort is not set"通常是指没有配置默认路由，网段 192.168.0.0/24、192.168.1.0/24 的路由是直连的，网段 192.168.2.0/24 的路由是通过 RIP 而得到的。

```
RouterA#show ip route
Gateway of last resort is not set

C    192.168.0.0/24 is directly connected, FastEthernet0/0
C    192.168.1.0/24 is directly connected, Serial2/0
R    192.168.2.0/24 [120/1] via 192.168.1.2, 00:00:15, Serial2/0
RouterA#
```

图 9.11　部分路由信息

9.2　开放最短路由优先协议

开放最短路径优先协议(Open Shortest Path First，OSPF)是通过对一个多路由器运行的网络设置多区域，区域间的路由是通过区域边界路由器(ABR)转发，区域内部的路由器只要知道到达 ABR 的路由就行了，这样就有效地减少了网络内部区域的路由条目。当整个网络经过链路状态信息的同步收敛，生成路由表后，OSPF 区域内部就可以直接转发数据。

9.2.1　OSPF 报文交换

OSPF 通过与直连路由器建立邻接关系互相传递链路状态信息，了解整个网络的拓扑结构，使得每个运行 OSPF 进程的路由器都装有一张"网络地图"。

1. OSPF 术语

为了理解 OSPF 的工作原理，下面给出一些运行 OSPF 以及与自治域路由相关的术语。

Hello 报文：多播方式发送，网络上的路由器通过 Hello 报文确定其相邻路由器并生成链路状态信息广播或通告(LSA)数据包。

自治系统(Autonomous System，AS)：在本章中，AS 是指路由器彼此相连，运行 OSPF 路由协议的所有路由器的集合。

邻居(neighbors)：同一个网段上的路由器都可以成为邻居。邻居是通过 Hello 报文来选择的，Hello 报文使用 IP 多播方式在每个端口定期发送。

邻接(adjacencies)：两台相邻路由器的双向关系。是邻居不一定就有相邻关系。成为邻接关系的路由器之间，不仅是进行 Hello 报文的交换，而且是要进行数据库的交换。

链路状态信息广播数据包(Link-State Advertisement，LSA)：它描述了指定链路内的路由信息，被传送给相邻路由器。

区域(area)：在运行 OSPF 的自治域内又划分为不同的区域，每个区域内都有一组路由

器，它们与其他路由器交换 LSA。每个区域都限制自己区域的 LSA，并要求汇总路由。

指定路由器(Designative Router，DR)：是一个运行 OSPF 的路由器，是被 OSPF Hello 协议推选的，通过它可以减少路由协议通信的数量和拓扑数据库的大小。在广播介质类型的网络中，网络中的每个路由器将自己的链路状态数据库向 DR 发送，而 DR 又将汇总的链路状态数据库向网络中的各个路由器广播。

备份指定路由器(Backup Designative Router，BDR)：备用 DR。当 DR 有问题时，由 BDR 接任其工作。

非指定路由器(DROTHER)：指在广播介质类型的网络中，除 DR、BDR 以外的所有路由器。

区域内路由器(Inter Area Router，IAR)：负责维护本区域内部路由器之间的链路状态数据库的路由器。

区域边界路由器(Area Border Router，ABR)：该路由器拥有所连接区域的所有链路状态数据库，并负责在区域之间发送 LSA 更新消息。

骨干(主干)路由器(backbone router)：自治域中连接主干链路的路由器，可以是区域内路由器，也可以是区域边界路由器。

自治系统边界路由器(Autonomous System Border Router，ASBR)：该路由器处于自治系统边界，负责和自治系统外部交换路由信息。

路由器 ID(Router ID，RID)：OSPF 协议使用一个被称为路由器 ID 的 32 位无符号整数来唯一标识一台路由器。基于这个目的，每一台运行 OSPF 的路由器都需要一个 RID。这个 RID 一般需要手工配置，可以将其配置为该路由器某个接口的 IP 地址。在没有手工配置 RID 的情况下，一些厂家的路由器支持自动从当前所有接口的 IP 地址选取一个 IP 地址作为 RID。

协议号(protocol numbers)：OSPF 采用 IP 报文直接封装 TCP 传输协议，协议号是 89。

邻居列表(neighbor list)：列出每台路由器全部已经建立邻接关系的邻居路由器。每个运行 OSPF 进程的路由器都要建立 3 张表，即邻居列表、链路状态数据库和路由表。

链路状态数据库(LSDB)：列出网络中所有路由器的信息，显示了全网的网络拓扑。

路由表(routing table)：OSPF 依据 Dijkstra 算法，从 LSDB 中计算得到一个以自己为树根的“最短路径树”，到最后每台路由器都将从最短路径树中构建自己的路由表。

2. OSPF 数据包

由于 OSPF 协议适用于大的区域，因而具有 5 种不同类型的数据包，见表 9.3。

表 9.3 OSPF 数据包类型

编　号	类　型	用　途
1	Hello	发现邻居，维持邻居关系，选举 DR/BDR
2	数据库描述	交换链路状态数据库 LSA 头
3	链路状态请求	请求一个指定的 LSA 数据细节
4	链路状态更新	发送被请求的 LSA 数据包
5	链路状态确认	对链路状态更新包的确认

OSPF 5 种不同类型的数据包的包头结构是一样的，数据包头部结构如图 9.12 所示，从包头的类型字段中就可以识别出各种数据包的类型，也就是对应于表 9.3 中的编号项(1～5)。以下介绍 5 种类型的 OSPF 数据包。

00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 位

版本	类型	总长度
路由器ID		
区域ID		
校验和		认证类型
身份认证		
身份认证		

图 9.12　OSPF 数据包头部结构

1）Hello 数据包

Hello 数据包是编号为 1 的 OSPF 数据包。运行 OSPF 协议的路由器每隔一定的时间发送一次 Hello 数据包，用以发现、保持邻居(neighbors)关系并可以选举 DR/BDR。

2）链路状态数据库描述数据包

链路状态数据库描述数据包(DataBase Description，DBD)是编号为 2 的 OSPF 数据包，该数据包在链路状态数据库交换期间产生。它的主要作用有 3 个：选举交换链路状态数据库过程中的主/从关系，确定交换链路状态数据库过程中的初始序列号和交换所有的 LSA 数据包头部。

3）链路状态请求数据包

链路状态请求数据包(LSA-REQ)是编号为 3 的 OSPF 数据包。该数据包用于请求在 DBD 交换过程发现的本路由器中没有的或已过时的 LSA 包细节。

4）链路状态更新数据包

链路状态更新数据包(LSA-Update)是编号为 4 的 OSPF 数据包。该数据包用于将多个 LSA 泛洪，也用于对接收到的链路状态更新进行应答。如果一个泛洪 LSA 没有被确认，它将每隔一段时间(默认是 5s)重传一次。

5）链路状态确认数据包

链路状态确认数据包(LSA-Acknowledgement)是编号为 5 的 OSPF 数据包。该数据包用于对接收到的 LSA 进行确认，会以组播的形式发送。如果发送确认的路由器状态是 DR 或者 BDR，确认数据包将被发送到 OSPF 路由器组播地址 224.0.0.5。如果发送确认的路由器的状态不是 DR 或者 BDR，确认将被发送到 OSPF 路由器组播地址 224.0.0.6。

3. 链路状态通告数据(LSA)

1）LSA 分类

OSPF 5 种类型数据包中的 4 种都封装了 LSA 或与 LSA 相关的信息，也可以理解为它们是在完成对 LSA 的操作。OSPF 是基于链路状态算法的路由协议，所有对路由信息的描述都是封装在 LSA 中发送出去的。LSA 头部格式如图 9.13 所示，LSA 根据不同的用途分为不同的类型(Type)，从包头的类型字段中就可以识别出各种 LSA 的类型。

00 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 位

LS年龄	选项	LS类型
链路状态ID		
通告路由器		
LS序列号		
LS校验和	长度	

图 9.13 LSA 头部格式

表 9.4 结合 LSA 头部格式，给出了目前使用最多的以上前 5 类 LSA 的类型、链路状态 ID、链路数据以及主要功能。

表 9.4 LSA 5 种类型及对应链路状态 ID 和链路数据

链路类型	链路状态 ID	链路数据	主要功能
Router LSA (Type1) (路由 LSA)	生成 LSA 的路由器 ID	IP 地址或标识	所有运行 OSPF 的路由器都会生成这种 LSA，主要描述本路由器运行 OSPF 的接口的连接、花费等
Network LSA (Type2) (网络 LSA)	该网络中 DR 的路由器 ID	IP 地址	本类型的 LSA 由 DR 生成。主要描述本网段中所有已经同其建立了邻接关系的路由器
Network Summary LSA(Type3) (网络汇总 LSA)	目标网络的 IP 地址	子网掩码	本类型的 LSA 也由 ABR 生成，描述某条路由的目的地址、掩码、花费值等信息
ASBR Summary LSA(Type4) (ASBR 汇总 LSA)	ASBR 路由器 ID	接口 IP 地址	本类型的 LSA 由 ABR 生成，而主要内容是描述到达本区域内部 ASBR 的路由
AS External LSA (Type5) (AS 外部 LSA)	目标网络的 IP 地址		本类型的 LSA 由 ASBR 生成，主要描述了到自治系统外部路由的信息

2）区域划分

在运行 OSPF 的自治域中，依据不同去向的 LSA，大部分路由器通常支持划分为多种区域：规则、存根、完全存根和准存根区域。

（1）规则区域，即不特别说明的区域，允许所有类型的 LSA 进出。路由器具有全部路由信息，有利选择到达目的地的最佳路径。不足之处是任何区域外的链路失效将引起局部的 OSPF 计算。

（2）存根区域(Stubby Area，STUB)，不允许外部的 LSA 进入。因此，ABR 不产生任何更新。外部 LSA 用于描述 OSPF 区域外的目的地，存根区域可以防止区域外部对本区域的影响，但并不能阻止 OSPF 区域内对该区域(area)的影响，仍然允许汇总 LSA，所以，其他区域将仍然影响到存根区域。

（3）完全存根(no-summary)区域同 STUB 类似，阻止外部 LSA。但是，与 STUB 不同的是，完全存根区域不允许汇总 LSA。这样其他区域将不会影响到完全存根区域。

（4）准存根区域(Not So Stubby Area，NSSA)与存根区域类似，但是，它可以将外部路

由导入到区域中。假如我们需要阻止外部 LSA 进入该区域,但域中的某个路由器(ASBR)仍然需要向区域外送外部 LSA,就需要使用 NSSA 区域。

9.2.2 OSPF 路由

1. 邻接建立过程

OSPF 邻接建立过程主要经过以下一些状态。

关闭(down)状态:没有发送 Hello 数据包,也没有收到 Hello 数据包。

尝试(attempt)状态:不停地向对方发送 Hello 数据包。当一个路由接口收到第一个 Hello 分组时,该路由器进入初始状态。

初始(init)状态:收到了对方的 Hello 数据包,但对方没有收到自己的 Hello 报文。

双向(two-way)状态:双方均收到了对方的 Hello 数据包。

启动(exstart)状态:发送 DBD 报文,选举主/从设备,设定初始序列号。

交换(exchange)状态:互相交换 LSA 报头信息。在交换状态,邻居路由器使用数据库描述数据包互相交换链路信息。如果路由器发现收到的数据包中有一条链路没有在自己的数据库中,则路由器将请求邻居路由器发送更新给自己。

装入(loading)状态:向对方请求自己没有的或过时的 LSA 信息,并在收到对方的更新 LSA 后添加到自己的链路状态数据库中。

完成(full)状态:双方的链路状态数据库完全相同。

2. OSPF 系统区域规划

随着网络规模的不断增大,网络拓扑结构也会不断发生变化,为了同步这种变化,网络中会有大量的 OSPF 协议报文在传递,这样必然会降低网络带宽的利用率。更糟糕的是,每一次变化都会导致网络中所有的路由器重新进行路由计算。解决这个问题的关键主要有两点:减少 LSA 的数量;屏蔽网络变化波及的范围。

OSPF 协议通过将自治系统划分成不同的区域(area)来解决上述问题。区域是在逻辑上将路由器划分为不同的组。区域的边界是路由器,这样会有一些路由器属于不同的区域,称作区域边界路由器(ABR),而一个网段只能属于一个区域。

每一个网段必须属于一个区域,或者说每个运行 OSPF 协议的接口必须指明属于某一个特定的区域,区域用区域号(Area ID)来标识。区域号是一个从 0 开始的 32 位整数。不同的区域之间可以通过 ABR 来传递路由信息。

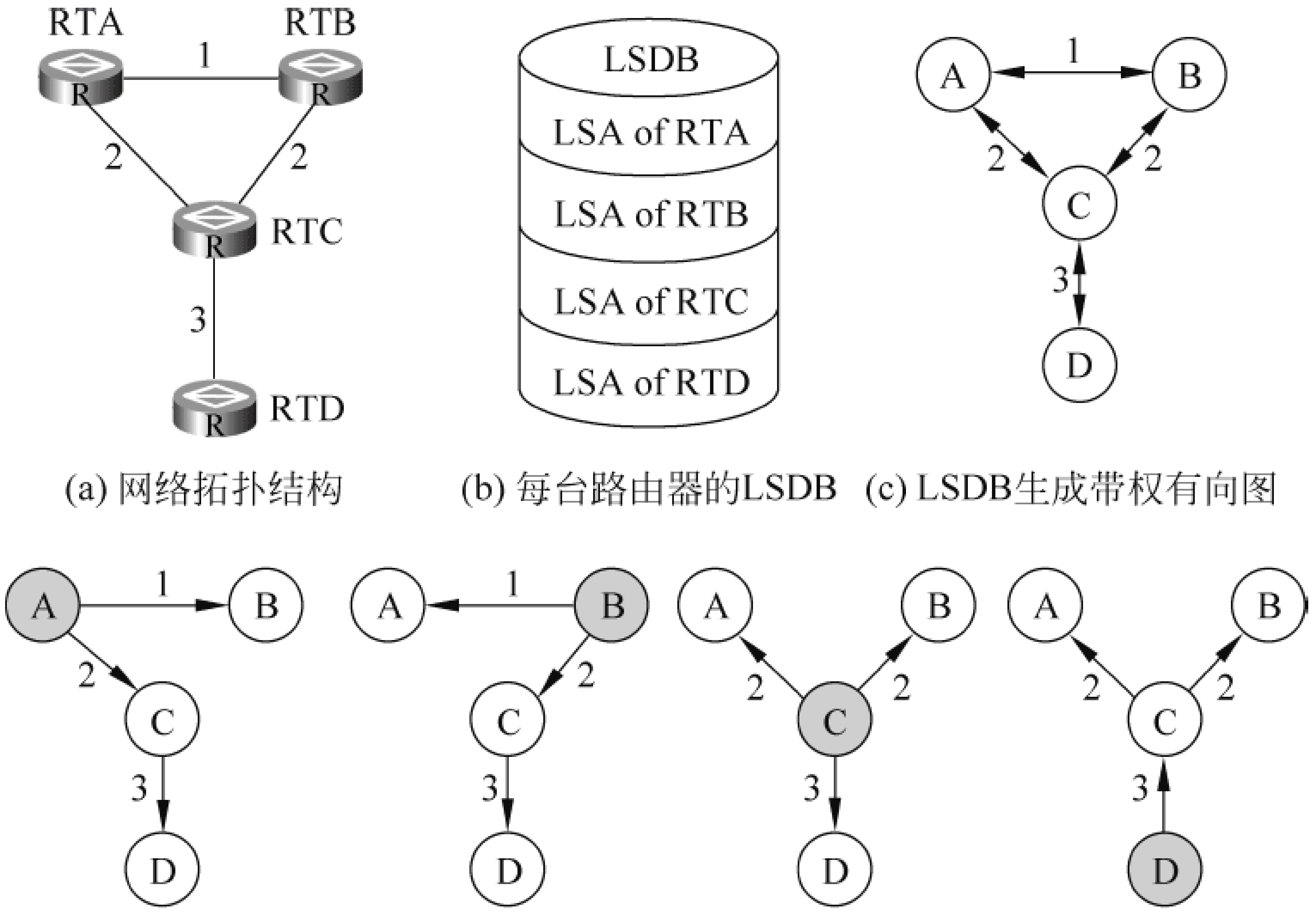
作为一个复杂的动态路由协议,在配置之前必须做好整个自治系统之内的规划。首先要选定的是合理地为 OSPF 划分区域,确定哪些路由器需要运行 OSPF 协议。

3. OSPF 计算过程

OSPF 协议计算出路由主要有以下几个步骤:描述本路由器周边的网络拓扑结构,并生成自己的 LSA,其中描述了自己的链路状态信息;将自己生成的 LSA 在自治系统中传播,并同时收集所有的其他路由器生成的 LSA,生成 LSDB;根据收集的所有的 LSA 计算路由,各路由器以自己的根节点计算出最小生成树,依据是链路的代价;将各路由器按照自己的最小生成树得出路由条目安装到路由表中。以下给出具体过程。

(1) 由 4 台路由器组成的网络,连线旁边的数字表示从一台路由器到另一台路由器所

需要的花费，如图 9.14(a)所示。这里，我们假定两台路由器相互之间发送报文所需花费是相同的。



(a) 网络拓扑结构　(b) 每台路由器的LSDB　(c) LSDB生成带权有向图

(d) 每台路由器以自己为根节点计算最小生成树

图 9.14　描述通过 OSPF 协议计算路由的过程

(2) 每台路由器都根据自己周围的网络拓扑结构生成一条 LSA(链路状态广播)，并通过相互之间发送协议报文将这条 LSA 发送给网络中其他的所有路由器。这样，每台路由器都收到了其他路由器的 LSA，所有的 LSA 放在一起称作 LSDB(链路状态数据库)。显然，4 台路由器的 LSDB 都是相同的，如图 9.14(b)所示。

(3) 由于一条 LSA 是对一台路由器周围网络拓扑结构的描述，那么 LSDB 则是对整个网络的拓扑结构的描述。路由器很容易将 LSDB 转换成一张带权的有向图，这张图便是对整个网络拓扑结构的真实反映。显然，4 台路由器得到的是一张完全相同的图，如图 9.14(c)所示。

(4) 接下来每台路由器在图中以自己为根节点，使用 SPF 算法计算出一棵最短路径树，由这棵树得到了到网络中各个节点的路由表。显然，4 台路由器各自得到的路由表是不同的。这样每台路由器都计算出了到其他路由器的路由，如图 9.14(d)所示。

4. 区域间路由计算

OSPF 将自治系统划分为不同的区域后，使得同一个区域内的路由器之间会保持 LSDB 的同步，网络拓扑结构的变化首先在区域内更新，而区域之间的路由计算则是通过 ABR 来完成的。

ABR 首先完成一个区域内的路由计算，然后查询路由表，为每一条 OSPF 路由生成一条 Type3 类型的 LSA，内容主要包括该条路由的目的地址、掩码、花费等信息。例如，如图 9.15 所示 Area ID=0，Type=3，DA=192.178.14.0，Mask=255.255.255.240，Metric=120 等，然后将这些 LSA 发送到另一个区域中。

在另一个区域中的路由器根据每一条 Type3 的 LSA 生成一条路由，由于这些路由信息都是由 ABR 发布的，所以这些路由的下一跳都指向该 ABR。

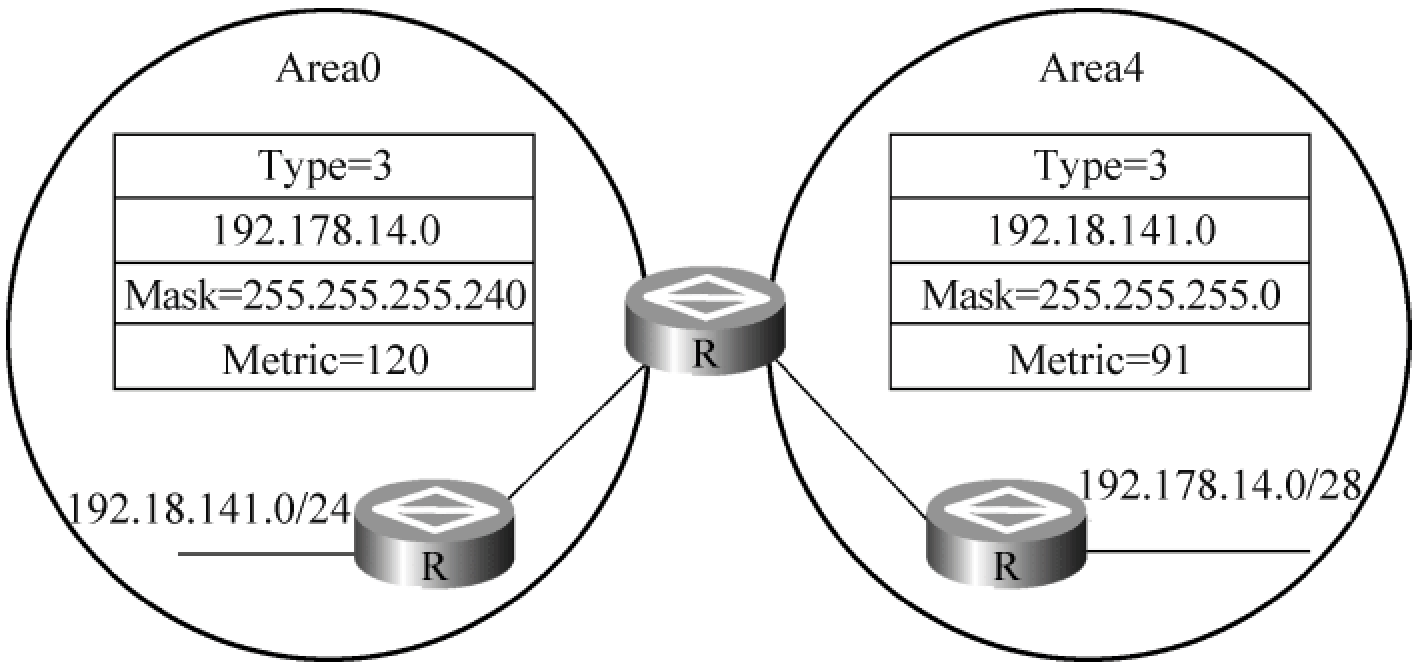


图 9.15 区域间的路由计算

5. 骨干区域与虚连接

自治系统被划分成一个以上的区域(area),就必须有一个区域是骨干区域(backbone area),并且保证其他区域与骨干区域直接相连或逻辑上相连,且骨干区域自身也必须是连通的。骨干区域是与众不同的,它的区域号(Area ID)是 0,即 Area0。

所有 ABR 将本区域内的路由信息封装成 LSA 后,统一发送给一个特定的区域,再由该区域将这些信息转发给其他区域。在这个特定区域内,每一条 LSA 都确切地知道生成者信息。在其他区域内所有到区域外的路由都会发送到这个特定区域中,这样就不会产生路由自环。这个“特定区域”就是骨干区域。

所有的区域必须和骨干区域相连,也就是说,每一个 ABR 连接的区域中至少有一个是骨干区域,而且骨干区域自身也必须是连通的。由于网络的拓扑结构复杂,有时无法满足每个区域必须和骨干区域直接相连的要求,例如图 9.16 中的 Area9。为解决此问题,OSPF 提出了虚连接的概念。虚连接是指在两台 ABR 之间,穿过一个非骨干区域(transit area,转换区域)建立的一条逻辑上的连接通道。可以理解为两台 ABR 之间存在一个点对点的连接。

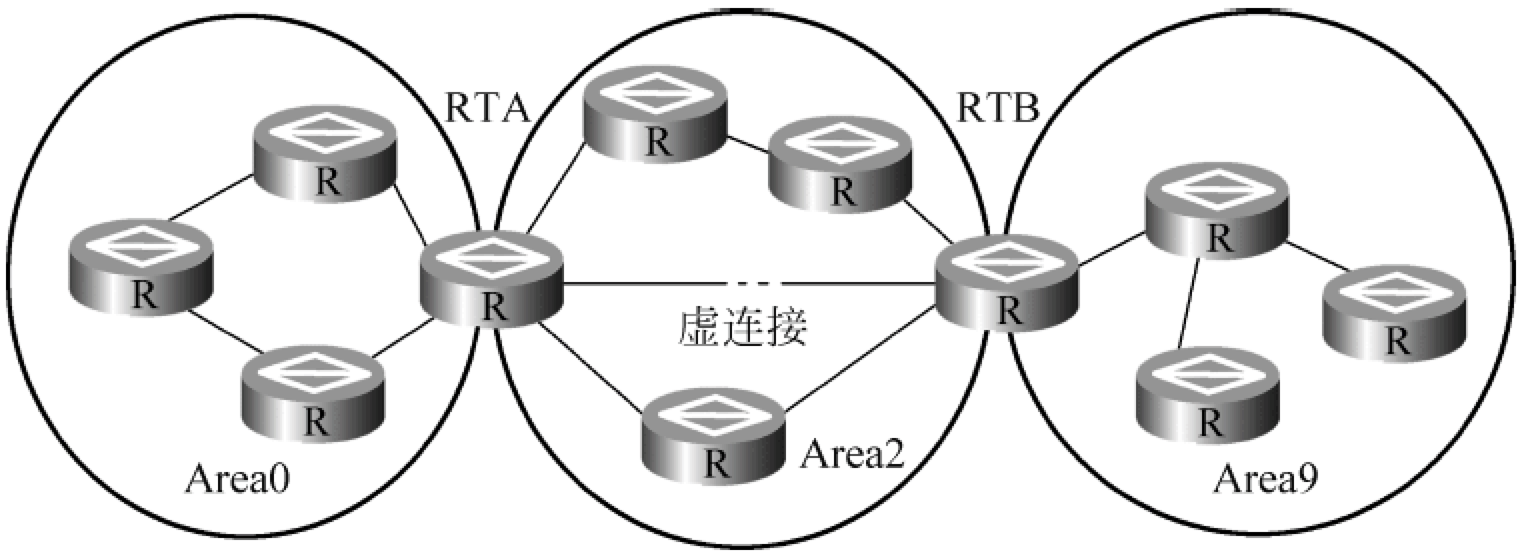


图 9.16 骨干区域与虚连接

逻辑通道是指两台 ABR 之间的多台运行 OSPF 的路由器只是起到一个转发报文的作用,这些报文对于这些路由器来说是透明的,只是当作普通的 IP 报文来转发。两台 ABR 之间直接传递路由信息。这里的路由信息是指由 ABR 生成的 Type3 的 LSA,区域内的路由器同步方式没有因此改变。

6. 区域与自治系统外部连接

1) 自治系统

OSPF 是自治系统(AS)内部路由协议,负责计算同一个自治系统内的路由。对于 OSPF 来说,整个网络只有“自治系统内”和“自治系统外”之分。需要注意的是:“自治系统外”并

不一定在物理上或拓扑结构中真正地位于自治系统的外部，而是指那些没有运行 OSPF 的路由器或者是某台运行 OSPF 协议的路由器中没有运行 OSPF 的接口。

2）自治系统边界路由器(ASBR)

作为一个内部协议的 IGP，OSPF 同样需要了解自治系统外部的路由信息，这些信息是通过 ASBR 获得的，自治系统边界路由器(ASBR)是那些将其他路由协议，也包括静态路由和接口的直接路由，发现的路由引入到 OSPF 中的路由器。ASBR 并不一定真的位于 AS 的边界，而是可以在自治系统中的任何物理位置。

3）计算自治系统外部路由

ASBR 为每一条引入的路由生成一条 Type5 的 LSA，主要内容包括该条路由的目的地址、掩码和花费等信息。这些路由信息将在整个自治系统中传播。计算路由时先在最短路径树中找到 ASBR 的位置，然后将所有由该 ASBR 生成的 Type5 类型的 LSA 都当作叶子节点挂在 ASBR 的下面。协议还规定：如果某个区域内有 ASBR，则这个区域的 ABR 在向其他区域生成路由信息时必须单独为这个 ASBR 生成一条 Type4 类型的 LSA，内容主要包括这个 ASBR 的 Router ID 和到它所需的花费值。

4）路由分级管理

OSPF 将所引入的自治系统外部路由分成两类：type1 和 type2。第一类是指引入 IGP 路由(如 RIP)，第二类外部路由是指引入的是 BGP 路由。

OSPF 一共将路由分为 4 级，按优先级从高到低排列：区域内路由→区域间路由→自治系统外一类路由(type 1)→自治系统外二类路由(type 2)。其中前两种路由在路由表中的优先级是一样的，默认值为 10；后两种路由在路由表中的优先级是相同的，默认值为 150。

9.2.3 单区域 OSPF 配置实例

单区域是指运行 OSPF 路由协议的路由器处于同一个区域(Area)。

单区域 OSPF 基本配置分为两个步骤：启动 OSPF 路由器协议进程和声明运行 OSPF 协议的路由器接口 IP 地址或子网地址。

1. 启动 OSPF 路由器协议进程

```
router ospf Process - ID
```

命令中，Process-ID 的范围为 1～65 535。该字段表示本地 OSPF 协议进程代号，只具有本地意义。在同一台路由器上运行多个 OSPF 协议实例时，OSPF 协议进程代号用于区别不同的 OSPF 协议进程。

2. 声明运行 OSPF 协议的路由器接口 IP 地址或子网地址

```
network A.B.C.D a.b.c.d area Area - id
```

命令中，A. B. C. D 代表网络地址号，a. b. c. d 代表 OSPF 通配符掩码；区域号：Area-id 的范围是 0～4 294 967 295。可以用两种格式表示：十进制数或 IP 地址的点分十进制形式。如，区域 1 可以表示为 1，也可以表示成 0.0.0.1。

【例 5】 要求完成同一区域内的路由器点到点链路 OSPF 单区域配置。

答：如图 9.17 所示的点到点链路中，路由器 RouterA、RouterB 通过串行接口互连。在 A、B 上各定义了一个环回接口 loopback 0，它们将成为各自路由器的 ID，并且用来模拟局

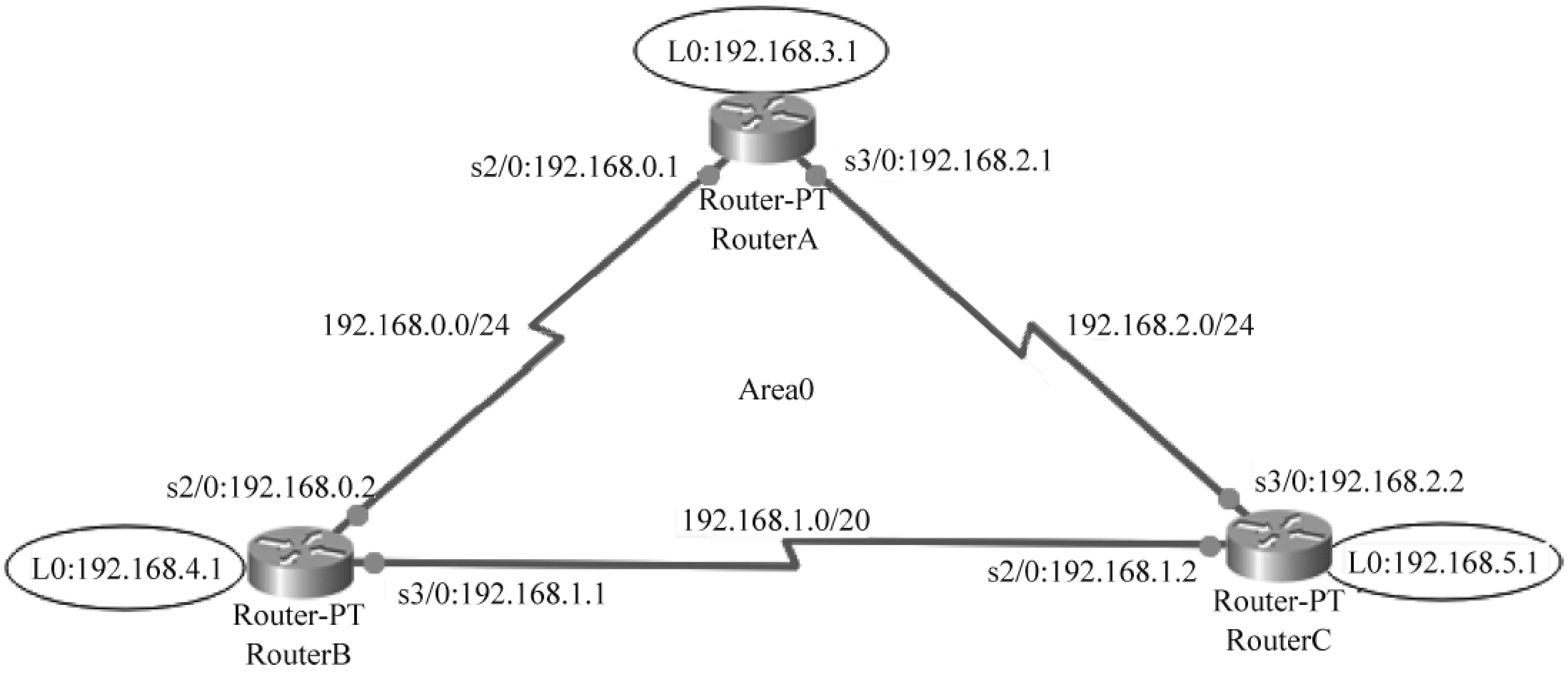


图 9.17　点到点链路 OSPF 配置

域网链路的逻辑接口。3 个路由器都处在同一区域：Area0。

（1）RouterA 的配置。

```
Router>enable
Router#configure terminal
Router(config)#hostname RouterA
RouterA(config-if)#interface loopback0
RouterA(config-if)#ip address 192.168.3.1 255.255.255.0
RouterA(config-if)#exit
RouterA(config)#router ospf 1                          !启动 OSPF,协议进程代号为 1
RouterA(config-router)#router-id 192.168.3.1           !配置路由器 ID 为 192.168.3.1
RouterA(config-router)#network 192.168.3.1 0.0.0.255 area 0   !设定区域 0 内网络
RouterA(config-router)#network 192.168.0.0 0.0.0.255 area 0   !0.0.0.255 为统配符掩码
RouterA(config-router)#network 192.168.2.0 0.0.0.255 area 0
RouterA(config-router)#exit
RouterA(config)#interface s2/0
RouterA(config-if)#ip address 192.168.0.1 255.255.255.0
RouterA(config-if)#clock rate 64000                           !s2/0 为 DCE
RouterA(config-if)#no shutdow
RouterA(config-if)#exit
RouterA(config)#interface s3/0
RouterA(config-if)#ip address 192.168.2.1 255.255.255.0
RouterA(config-if)#clock rate 64000                           !s3/0 为 DCE
RouterA(config-if)#no shutdow
RouterA(config-if)#exit
RouterA(config)#
```

（2）RouterB 的配置。

RouterB 的具体配置参考截图 9.18，其中 s3/0 为 DCE。

（3）RouterC 的配置。

RouterC 的具体配置，参看图 9.19。

```
Router>enable
Router#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
Router(config)#hostname RouterB
RouterB(config)#interface loopback0
RouterB(config-if)#
%LINK-5-CHANGED: Interface Loopback0, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface Loopback0, changed state
to up
RouterB(config-if)#ip address 192.168.4.1 255.255.255.0
RouterB(config-if)#exit
RouterB(config)#router ospf 2
RouterB(config-router)#router-id 192.168.4.1
RouterB(config-router)#network 192.168.4.1 0.0.0.255 area 0
RouterB(config-router)#network 192.168.0.0 0.0.0.255 area 0
RouterB(config-router)#network 192.168.1.0 0.0.0.255 area 0
RouterB(config-router)#exit
RouterB(config)#interface s2/0
RouterB(config-if)#ip address 192.168.0.2 255.255.255.0
RouterB(config-if)#no shutdow
RouterB(config-if)#
%LINK-5-CHANGED: Interface Serial2/0, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface Serial2/0, changed state
to up
RouterB(config-if)#exit
RouterB(config)#interface s3/0
RouterB(config-if)#ip address 192.168.1.1 255.255.255.0
RouterB(config-if)#clock rate 64000
RouterB(config-if)#no shutdow
%LINK-5-CHANGED: Interface Serial3/0, changed state to down
RouterB(config-if)#exit
RouterB(config)#
```

图 9.18 路由器 B 的配置截图

```
Router>enable
Router#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
Router(config)#hostname RouterC
RouterC(config)#interface loopback0
RouterC(config-if)#
%LINK-5-CHANGED: Interface Loopback0, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface Loopback0, changed state
to up
RouterC(config-if)#ip address 192.168.4.1 255.255.255.0
RouterC(config-if)#exit
RouterC(config-if)#router ospf 3
RouterC(config-router)#router-id 192.168.5.1
RouterC(config-router)#network 192.168.5.1 0.0.0.255 area 0
RouterC(config-router)#network 192.168.2.2 0.0.0.255 area 0
RouterC(config-router)#network 192.168.1.2 0.0.0.255 area 0
RouterC(config-router)#exit
RouterC(config)#interface s2/0
RouterC(config-if)#ip address 192.168.1.2 255.255.255.0
RouterC(config-if)#no shutdow
RouterC(config-if)#
%LINK-5-CHANGED: Interface Serial2/0, changed state to up
RouterC(config-if)#exit
RouterC(config)#interface s3/0
RouterC(config-if)#ip address 192.168.2.2 255.255.255.0
RouterC(config-if)#no shutdow
%LINK-5-CHANGED: Interface Serial3/0, changed state to down
RouterC(config-if)#exit
RouterC(config)#
```

图 9.19 路由器 C 的配置截图

(4) 诊断及检查。

为了方便读者课后动手练习,下面给出一些与 OSPF 相关的诊断及检查的命令。

```
Show running - config            ! 显示运行配置
Show ip route                    ! 显示 IP 路由信息
Show ip route ospf               ! 仅显示 OSPF 路由
Show ip ospf process - id        ! 显示与特定进程 ID 相关的信息
```

```
Show ip ospf                       !显示 OSPF 相关信息
Show ip ospf border-routers        !显示边界路由器
Show ip ospf database              !显示 OSPF 的归纳数据库
Show ip ospf interface             !显示指定接口上的 OSPF 信息
Show ip ospf neighbor              !显示 OSPF 相邻信息
Show ip ospf request-list          !显示链路状态请求列表
Show ip ospf summary-address       !显示归纳路由的再发布信息
Show ip ospf virtual-links         !显示虚拟链路信息
Show ip interface                  !显示接口的 IP 设置
debug ip ospf adj                  !调试或查看有哪些邻接关系及邻接双方接口信息
debug ip ospf events               !调试或查看 OSPF 事件
debug ip ospf flood                !调试或查看 OSPF 泛洪
debug ip ospf packet               !调试或查看 OSPF 报文
debug ip ospf spf                  !调试或查看 OSPF 的 SPF
```

例如，在路由器 RouterA 上执行 Show ip route 后显示的信息如图 9.20 所示，它是路由器 RouterA 的路由信息。通过执行 Show ip interface 命令，路由器各端口的信息就会一目了然，路由器 RouterA 端口如图 9.21 所示。如在 PC 上通过执行 ping 命令，显示各个设备之间能够正常收发数据包，说明配置正确。

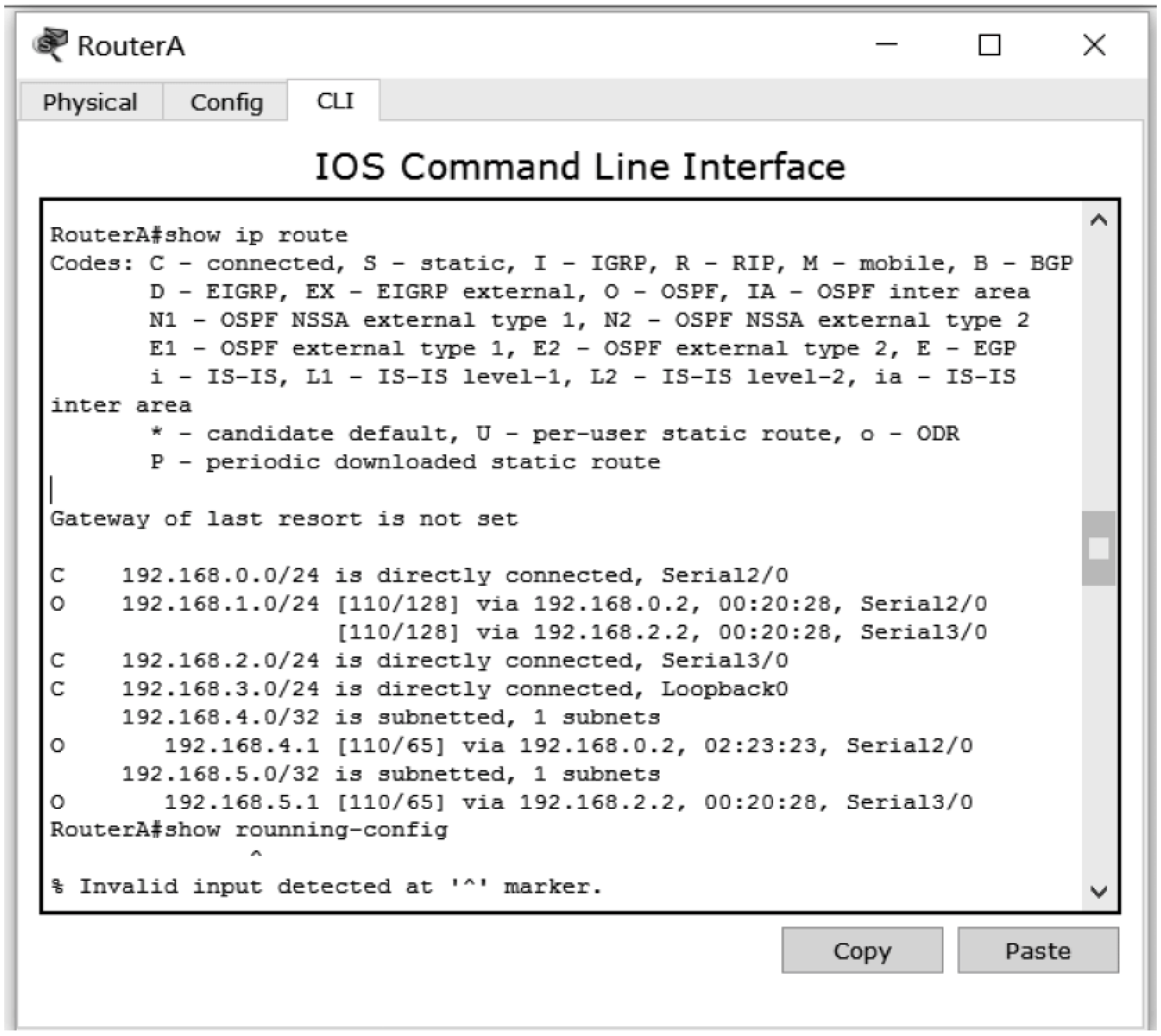


图 9.20 RouterA 显示路由截图

```
Port              Link   IP Address        IPv6 Address                         MAC Address
FastEthernet0/0   Down   <not set>         <not set>                            0010.1145.874E
FastEthernet1/0   Down   <not set>         <not set>                            0060.2FD8.0633
Serial2/0         Up     192.168.0.1/24    <not set>                            <not set>
Serial3/0         Up     192.168.2.1/24    <not set>                            <not set>
FastEthernet4/0   Down   <not set>         <not set>                            0040.0B4C.3C0C
FastEthernet5/0   Down   <not set>         <not set>                            000B.BE60.3425
Loopback0         Up     192.168.3.1/24    <not set>                            00D0.BCD9.0462
Hostname: RouterA
```

图 9.21 RouterA 端口信息截图

9.2.4 多区域 OSPF 配置实例

多区域是指运行 OSPF 路由协议的路由器可能处在不同的区域(area)。

1. 多区域 OSPF 配置

【例 6】 多区域 OSPF 配置如图 9.22 所示,它有两个区域 Area0 和 Area1。在配置中,需要区分清楚配置的网段在哪个区域。路由器 RouterA、RouterB 和 RouterD、RouterE 都是分别在单独的区域中,配置时同属一个区域的路由器各端口区域号相同,而路由器 RouterB 同时在两个区域中,配置时不同的网段声明的区域也不相同。具体配置如下。

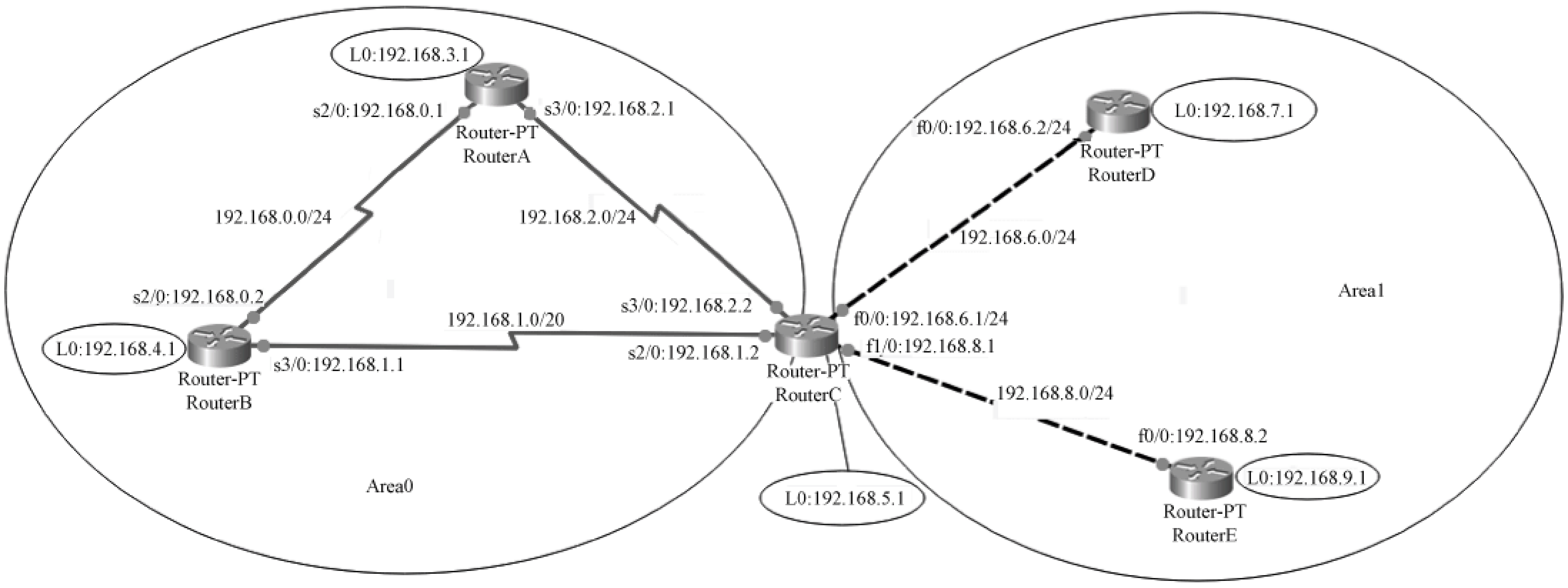


图 9.22 多区域 OSPF 配置

答:(1) RouterA 的配置。

由于 Area0 与例 5 题的图 9.17 相同,RouterA 属于区域内路由器(IAR),它与例 5 的 RouterA 设置也完全一样,只是增加一条以下默认路由即可。

```
RouterA(config)# ip route 0.0.0.0 0.0.0.0 192.168.2.2
```

(2) RouterB 的配置。

RouterB 属于 IAR,与例 5 的 RouterB 设置也完全一样,增加以下默认路由即可。

```
RouterB(config)# ip route 0.0.0.0 0.0.0.0 192.168.1.2
```

(3) RouterC 的配置。

由于 RouterC 是区域边界路由器(ABR),同属于 Area0、Area1,并连接两个区域,所以 RouterC 在保持例 5 原配置的基础上,增加以下路由配置即可。

```
RouterC> enable
RouterC# configure terminal
RouterC(config)# router ospf 1
RouterC(config-router)# network 192.168.8.0 0.0.0.255 area 1
RouterC(config-router)# network 192.168.6.0 0.0.0.255 area 1
RouterC(config-router)# exit
RouterC(config)# interface FastEthernet1/0
RouterC(config-if)# ip address 192.168.8.1 255.255.255.0
RouterC(config-if)# no shutdown
RouterC(config-if)# exit
```

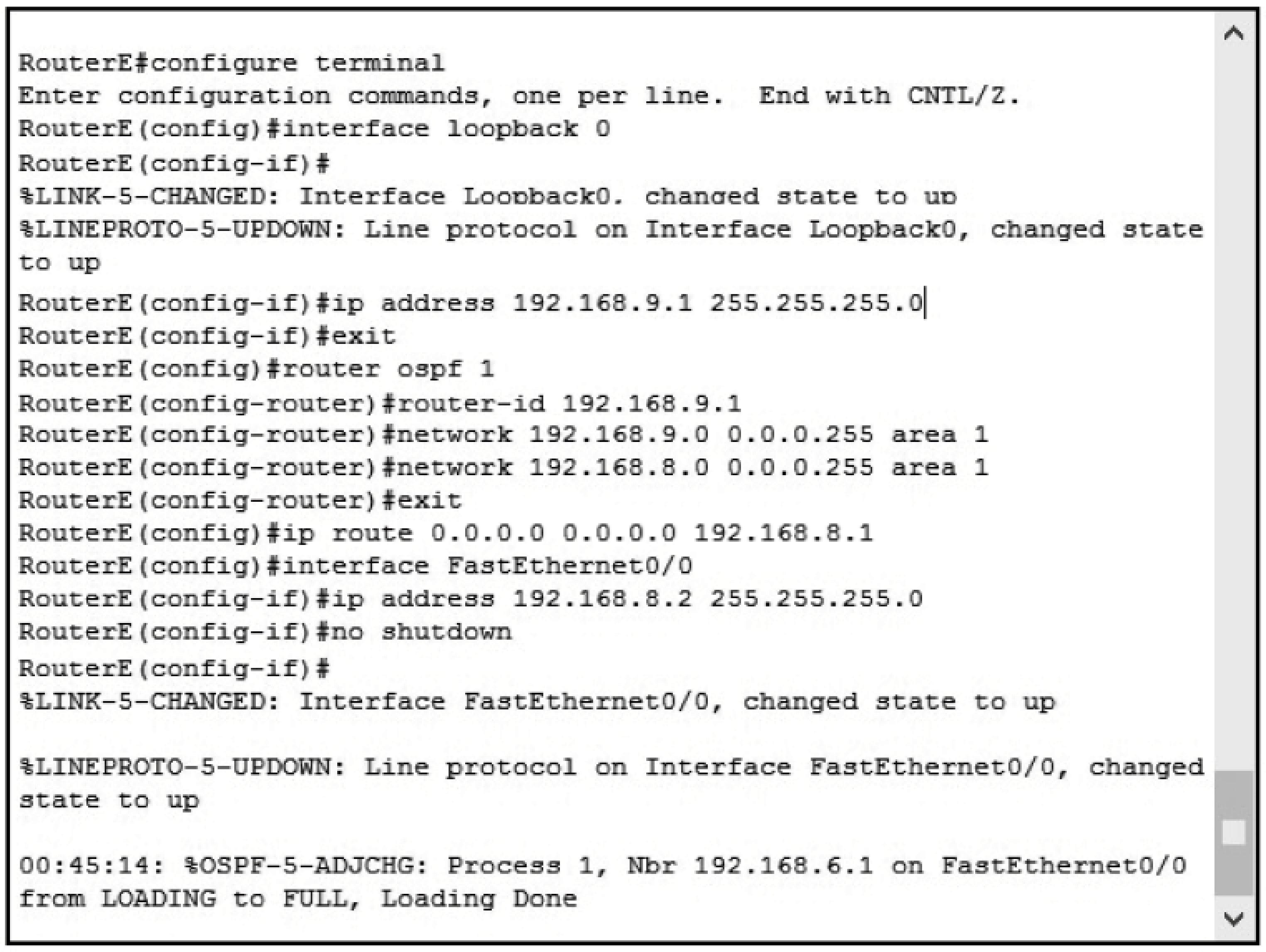
```
RouterC(config)#interface FastEthernet0/0
RouterC(config-if)#ip address 192.168.6.1 255.255.255.0
RouterC(config-if)#no shutdown
```

(4) RouterD 的配置。

```
RouterD(config)#interface loopback 0
RouterD(config-if)#ip address 192.168.7.1 255.255.255.0
RouterD(config-if)#exit
RouterD(config)#router ospf 1
RouterD(config-router)#router-id 192.168.7.1
RouterD(config-router)#network 192.168.7.0 0.0.0.255 area 1
RouterD(config-router)#network 192.168.6.0 0.0.0.255 area 1
RouterD(config-router)#exit
RouterD(config)#ip route 0.0.0.0 0.0.0.0 192.168.6.1
RouterD(config)#interface FastEthernet 0/0
RouterD(config-if)#ip address 192.168.6.2 255.255.255.0
RouterD(config-if)#no shutdown
```

(5) RouterE 的配置。

RouterE 配置如图 9.23 所示。



```
RouterE#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
RouterE(config)#interface loopback 0
RouterE(config-if)#
%LINK-5-CHANGED: Interface Loopback0, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface Loopback0, changed state
to up
RouterE(config-if)#ip address 192.168.9.1 255.255.255.0
RouterE(config-if)#exit
RouterE(config)#router ospf 1
RouterE(config-router)#router-id 192.168.9.1
RouterE(config-router)#network 192.168.9.0 0.0.0.255 area 1
RouterE(config-router)#network 192.168.8.0 0.0.0.255 area 1
RouterE(config-router)#exit
RouterE(config)#ip route 0.0.0.0 0.0.0.0 192.168.8.1
RouterE(config)#interface FastEthernet0/0
RouterE(config-if)#ip address 192.168.8.2 255.255.255.0
RouterE(config-if)#no shutdown
RouterE(config-if)#
%LINK-5-CHANGED: Interface FastEthernet0/0, changed state to up

%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/0, changed
state to up

00:45:14: %OSPF-5-ADJCHG: Process 1, Nbr 192.168.6.1 on FastEthernet0/0
from LOADING to FULL, Loading Done
```

图 9.23　RouterE 配置

2. OSPF 配置检查

各个路由器配置完成后，用 Show ip route 命令查看每个路由表的信息，以便检查路由配置信息。图 9.24 是 RouterC 的路由信息，里面有直连(C)路由，也有来自 OSPF(O)的动态路由，Area0、Area1 的所有路由在这里都能找到。

图 9.25 是 RouterE 的路由信息，除了直连路由和动态路由以外，还有一条静态(S)路由，并且只有 Area1 的路由信息，体现了通过分小区(area)减小了路由表的存储压力，并通过静态配置的默认路由将离开 Area1 的数据包送到区域边界路由器 RouterC 处理。

```
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS
inter area
       * - candidate default, U - per-user static route, o - ODR
       P - periodic downloaded static route

Gateway of last resort is not set

O    192.168.0.0/24 [110/128] via 192.168.1.1, 00:37:37, Serial2/0
                    [110/128] via 192.168.2.1, 00:37:37, Serial3/0
C    192.168.1.0/24 is directly connected, Serial2/0
C    192.168.2.0/24 is directly connected, Serial3/0
     192.168.3.0/32 is subnetted, 1 subnets
O       192.168.3.1 [110/65] via 192.168.2.1, 00:37:37, Serial3/0
     192.168.4.0/32 is subnetted, 1 subnets
O       192.168.4.1 [110/65] via 192.168.1.1, 00:37:37, Serial2/0
C    192.168.5.0/24 is directly connected, Loopback0
C    192.168.6.0/24 is directly connected, FastEthernet0/0
     192.168.7.0/32 is subnetted, 1 subnets
O       192.168.7.1 [110/2] via 192.168.6.2, 00:37:02, FastEthernet0/0
C    192.168.8.0/24 is directly connected, FastEthernet1/0
     192.168.9.0/32 is subnetted, 1 subnets
```

图 9.24　RouterC 的路由信息

```
O    192.168.6.0/24 [110/2] via 192.168.8.1, 00:26:07, FastEthernet0/0
     192.168.7.0/32 is subnetted, 1 subnets
O       192.168.7.1 [110/3] via 192.168.8.1, 00:26:07, FastEthernet0/0
C    192.168.8.0/24 is directly connected, FastEthernet0/0
C    192.168.9.0/24 is directly connected, Loopback0
S*   0.0.0.0/0 [1/0] via 192.168.8.1
RouterE#
```

图 9.25　RouterE 的路由信息

9.2.5　OSPF 其他配置

协议中的命令是一台路由器运行 OSPF 协议所必需的，还有一些命令虽然不是必须配置的，但如果配置得当，就可以减少网络中的路由信息量和路由表规模等。以下将给出针对具体不同区域的 OSPF 相关命令配置，并通过拓扑图 9.26 进行配置举例。

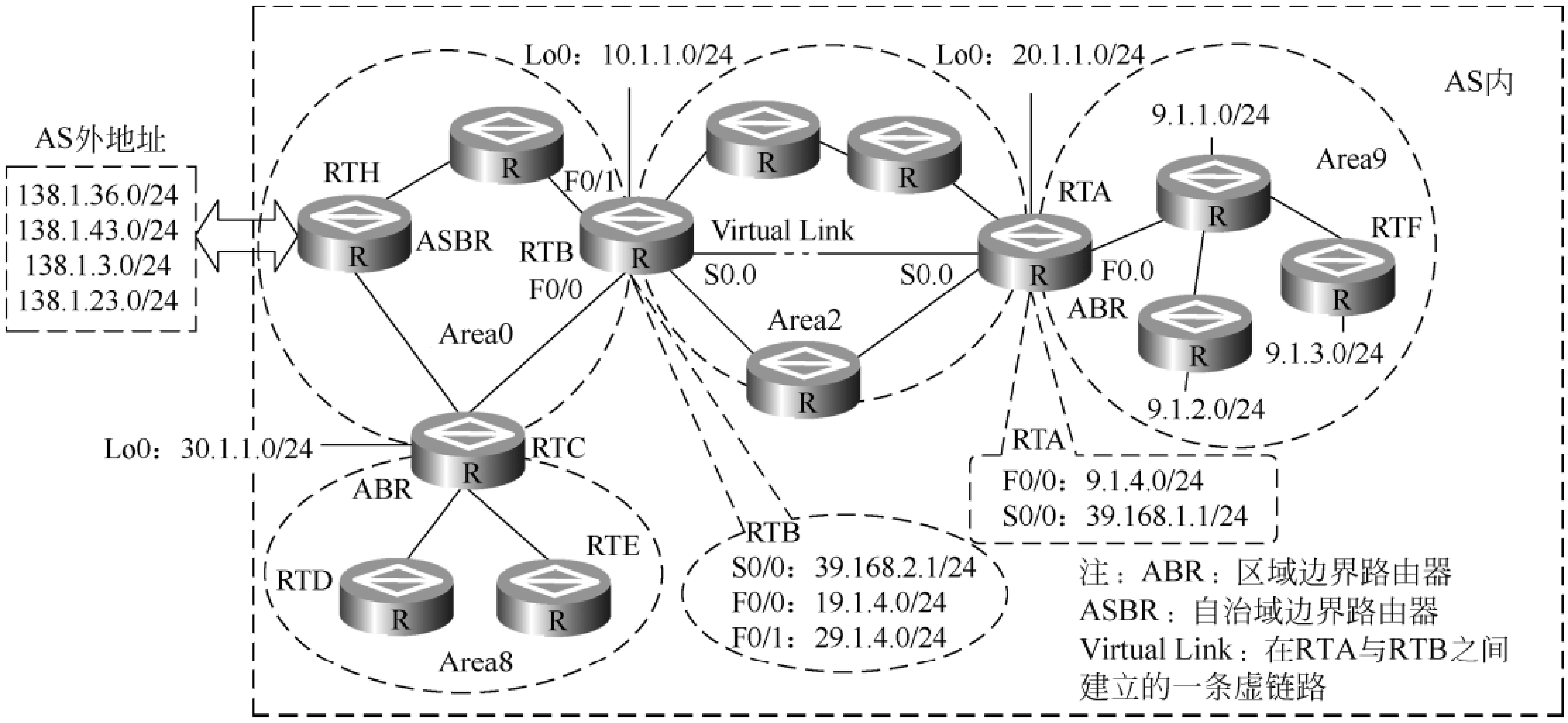


图 9.26　STUB 区域、路由聚合和虚连接

1. 配置区域间路由聚合

(1) 在 ASBR 上汇总配置。

```
summary - address ip_address mask_address
```

ip_address 为汇总 IP 地址,mask_address 为掩码。

【例 7】 图 9.26 中,ASBR(边界路由器)用汇总路由 138.1.0.0/16 代替被覆盖的 AS 外部网段 138.1.36.0/24,138.1.43.0/24,138.1.3.0/24,138.1.23.0/24。并以 5 类 LSA 的形式,将汇总路由在整个 AS 区域内传播。

答: 在 Area0 的 RTH 上配置路由聚合命令如下。

```
RouterH(config - router) # summary - address 138.1.0.0 255.255.0.0
```

(2) 在 ARB 上汇总配置。

```
area area_id range summary - address mask_address
```

area_id 指定了区域 ID,把 OSPF 区域内的多条路由合并成一条,summary-address 为合并后的 IP 地址,与后面的 mask_address 配合使用。

【例 8】 图 9.26 中,Area9 中有 3 条路由 9.1.1.0/24,9.1.2.0/24,9.1.3.0/24。希望默认这 3 条路由配置时,也会被发送到其他区域。如果在 Area9,在担任区域边界路由器(ABR)的 RTA 上配置路由聚合,则可以将这 3 条路由变为 1 条。

答: 其配置命令如下。

```
RouterA(config - router) # area 9 range 9.1.0.0 255.255.0.0
```

2. 配置存根区域和完全存根区域

```
area area - id stub [no - summary]
```

此命令可以将一个区域定义成存根区域(STUB)或完全存根区域。其中,no-summary 用于将一个区域定义为完全存根区域,该参数只在区域边界路由器上使用。

再看图 9.26 的 Area0,有一台 ASBR,引入了 4 条自治系统外的路由,默认这 4 条路由的将被发送到整个自治系统中。在实际的运行情况下,有时自治系统中的大部分路由都是这种自治系统外部路由。由于在存根区域内,路由器的链路状态数据库不含有 5 类和 4 类 LSA,所以为了减小路由表的规模,协议规定可以将一部分区域规定为 STUB 区域,在这种区域中是不会传播自治系统外部路由的,但是为了保证路由可达,由该区域的 ABR 生成一条默认路由发布到 STUB 区域内。

【例 9】 将图 9.26 中 Area8 指定为 STUB 区域,假设 Area8 符合 STUB 区域的条件。

答:

```
RouterC(config - router) # area 8 stub
```

以上是 RTC 的配置。因为一个区域要配置成 STUB 区域,区域内所有的路由器都需要配置该属性,所以在 RTC 和 RTE 上同样需要配置 area 8 stub 命令。

STUB 区域也有一个缺陷,就是 STUB 区域内不能存在 ASBR,即图中的 RTD 和 RTE 都不能再引入其他路由协议发现的路由。

3. 配置准存根区域

```
area area_id nssa [ default - route - advertise ] [ no - import - route ] [ no - summary ]
```

area_id 是需要配置成 NSSA 区域的区域号。[ ]内的参数只有在该路由器是 ABR 时才会生效。即,如果路由器只是一台区域内路由器,只需要配置 area area_id nssa。

default-route-advertise:配置该参数后,ABR 会向 NSSA 内部发送一条默认路由。

no-summary:配置该参数后,成为完全准存根区域,ABR 会将 Type3 类型的 LSA 也过滤掉,即:NSSA 区域中也不会出现区域间路由,路由表进一步精简。

no-import-route:配置该参数后,ABR 自身引入的外部路由,不再以 Type7 类型的 LSA 的形式在 NSSA 区域中传递。该参数推荐配置。

【例 10】 将图 9.26 中的 Area9 指定为 NSSA 区域。

答:

```
RouterF(config - router)#area 9 nssa                !在非边界路由器定义,其他定义类似于 RTF
```

【例 11】 将图 9.26 中的 Area9 指定为完全 NSSA 区域。

答:

```
RouterA(config - router)#area 9 nssa no - summary  !在边界路由器 RTA 上定义
```

4. 配置 OSPF 邻居认证

1) 相邻路由器明文认证的设置

启动明文认证的命令格式如下:

```
Area area_id authentication
```

在接口上配置明文认证和密码的命令格式如下:

```
ip ospf authentication - key password
```

其中,password 为明文认证字(密码),由 1~8 个字符构成。

【例 12】 配置图 9.26 中 Area0,RouterB 的 fastethernet 0/1 对 OSPF 报文采用明文认证,密码为 admin。

答: 命令配置如下。

```
Router#hostname RouterB
RouterB(config)#router ospf 1
RouterB(config - router)#area 0 authentication
RouterB(config - router)#exit
RouterB(config)#interface fastethernet 0/1
RouterB(config - if)#ip ospf authentication - key admin
```

2) 相邻路由器 MD5 密文认证的设置

启动区域 MD5 密文认证的命令格式如下:

```
Area area_id authentication message - digest
```

在接口上配置 MD5 密文认证和认证字键值的命令格式如下:

```
ip ospf message - digest - key key - id md5 key
```

其中,key-id 为 MD5 认证方式时的认证字键值,是一个整数,取值范围为 1～255。key 为 MD5 认证加密的密钥,是一个字符串,由 1～16 个字符构成。

【例 13】 配置图 9.26 中 Area0 内,RouterB 的 fastethernet 0/0 对 OSPF 报文采用 MD5 密文认证,MD5 认证密码密钥为 admin,key-id 为 6。

答:命令配置如下。

```
Router#hostname RouterB
RouterB(config)#router ospf 1
RouterB(config - router)#area 0 authentication message - digest
RouterB(config - router)#exit
RouterB(config)#interface fastethernet 0/1
RouterB(config - if)#ip ospf message - digest - key 6 md5 admin
```

需要说明的是,如对某条链路进行认证,其相邻(相连)的两个路由器都必须要配置。

5. 配置虚连接

根据协议规定,所有的非骨干区域必须与骨干区域(Area0)连通。当一个区域无法与骨干区域直接连通时(如图 9.26 中的 Area9),需要配置一条虚连接。虚连接是在两台 ABR 之间配置,中间穿过一个非骨干区域(transit-area)。其配置命令如下:

```
area area - id virtual - link router - id [hello - interval seconds] [retransmit - interval
seconds] [retransmit - delay seconds] [dead - interval seconds] [[authentication - key key ] |
[ message - digest - key key - id md5 key ]]
```

area-id:传输区域 ID 号。

router-id:形成虚链路的对端路由器 ID 号。

hello-interval:配置虚连接的 Hello 报文发送间隔,单位为 s,默认值为 10s。

retransmit-interval:配置虚连接的重传间隔,单位为 s,默认值为 5s。

retransmit-delay:配置虚连接的 LSA 的传输延迟,单位为 s,默认值为 1s。

dead- interval:配置虚连接的邻接点死亡时间,单位为 s,默认值为 40s。

authentication-key:邻居路由器明文验证方式关键字。

message-digest-key key-id md5 key:邻居路由器 MD5 密文验证方式关键字。

在配置虚连接时,两端都必须要配置。

【例 14】 图 9.26 中,需要在 RTA(router id = 20.1.1.0)和 RTB(router id = 10.1.1.0)之间配置虚连接,穿过 Area2。要求对路由器 RTA、RTB 进行详细配置。

答:(1) 对 RTA 的命令配置。

```
Router>enable
Router#config terminal
Router#hostname RouterA
RouterA(config)#interface loopback0                !配置 loopback 接口,用于 router - id
RouterA(config - if)#ip address 20.1.1.0 255.255.255.0
RouterA(config - if)#no shutdown
RouterA(config - if)#interface fa0/0
RouterA(config - if)#ip address 9.1.4.0 255.255.255.0
```

```
RouterA(config-if)#no shutdown
RouterA(config-if)#interface s0/0
RouterA(config-if)#ip address 39.168.1.1 255.255.255.0
RouterA(config-if)#no shutdown
RouterA(config-if)#clockrate 64000              !连接该接口的电缆若属于 DTE,则不用配置
RouterA(config-if)#exit
RouterA(config)#router ospf 1
RouterA(config-router)#network 9.1.4.0 0.0.0.255 area9
RouterA(config-router)#network 39.168.1.1 0.0.0.255 area2
RouterA(config-router)#area 2 virtual-link 10.1.1.0          !需指定对端(RTB)的 router id
RouterA(config-router)#end
RouterA#
```

(2) 对 RTB 的命令配置。

```
Router>enable
Router#config terminal
Router#hostname RouterB
RouterB(config)#interface loopback0              !配置 loopback 接口,用于 router-id
RouterB(config-if)#ip address 10.1.1.0 255.255.255.0
RouterB(config-if)#no shutdown
RouterB(config-if)#interface fa0/0
RouterB(config-if)#ip address 19.1.4.0 255.255.255.0
RouterB(config-if)#no shutdown
RouterB(config-if)#interface fa0/1
RouterB(config-if)#ip address 29.1.4.0 255.255.255.0
RouterB(config-if)#no shutdown
RouterB(config-if)#interface s0/0
RouterB(config-if)#ip address 39.168.2.1 255.255.255.0
RouterB(config-if)#no shutdown
RouterB(config-if)#clockrate 64000              !连接该接口的电缆若属于 DTE,则不用配置
RouterB(config-if)#exit
RouterB(config)#router ospf 1
RouterB(config-router)#network 19.1.4.0 0.0.0.255 area0
RouterB(config-router)#network 29.1.4.0 0.0.0.255 area0
RouterB(config-router)#network 39.168.2.1 0.0.0.255 area2
RouterB(config-router)#area 2 virtual-link 20.1.1.0          !需指定对端(RTA)的 router id
RouterB(config-router)#end
RouterB#
```

9.3 中间系统到中间系统协议

9.3.1 IS-IS 工作原理

1. IS-IS 路由器分类

根据路由器所处的网络位置和区域类型不同,中间系统到中间系统(Intermediate System to Intermediate System,IS-IS)将路由器分为三类:Level-1、Level-2 和 Level-1-2。

Level-1 路由器是普通区域中的路由器,只能存在于非骨干区域中。并且,Level-1 路由器只能与所处同一区域的 Level-1 或 Level-1-2 路由器建立 Level-1 邻居关系,交换路由信息。在工作模式中,Level-1 路由器只能转发同一区域内的路由报文,而通往外部区域的路

由信息将被转发至最近的 Level-1-2 路由器，再由其进行转发。

Level-2 路由器是骨干区域中的路由器，与其他的 Level-2 路由器处于同一骨干区域。且 Level-2 路由器只能同所处同一骨干区域的 Level-2 路由器，或是其他区域的 Level-1-2 路由器建立 Level-2 邻居关系。在 IS-IS 网络中，所有的 Level-2 与 Level-1-2 路由器连接在一起，共同组成一个骨干网络，也称为 Level-2 区域。与骨干区域不同，Level-2 区域是连接了整个 IS-IS 网络内，转发 Level-2 路由的设备集合。

Level-1-2 路由器用于骨干区域与非骨干区域的区域间连接。Level-1-2 路由器既可以与 Level-1 路由器建立邻居关系，进行 Level-1 路由的转发，也可以与 Level-2 路由器建立邻居关系，转发 Level-2 路由。同时，Level-1-2 路由器不一定必须在区域边界上，也可以存在于骨干区域或非骨干区域的内部。

2. IS-IS 报文交换

IS-IS 作为一种链路状态路由协议，在 IS-IS 网络中，每一台路由器都会生成自己的链路状态报文（Link State Packet，LSP），它包含了路由器所有的 IS-IS 协议接口的链路状态信息。通过与相邻设备建立 IS-IS 邻居关系，互相更新本地设备的链路状态数据库（Link State DataBase，LSDB）。更新过程使得整个 IS-IS 网络中的各设备的 LSDB 实现同步，并根据 SPF 算法计算出 IS-IS 路由，加入到设备的路由表中，指导报文转发。

IS-IS 网络中，链路信息的传递通过 PDU（协议数据单元）来完成，其使用的 PDU 类型主要有 3 种：Hello PDU，LSP 和序列号 PDU（Sequence Number PDU，SNP）。

Hello PDU 报文用于建立和维持 IS-IS 网络中的邻居关系，由设备周期性地发送给自己的 IS-IS 邻居设备。其中，Level-1 邻居和 Level-2 邻居所发送的 Hello PDU 报文并不相同，类型号分别为 15 和 16。在非广播网中则使用类型号为 17 的 Hello PDU。

LSP 是包含 IS-IS 路由链路状态信息的 PDU 报文，用于和其他设备交换链路状态信息。每一台设备都会产生自己的 LSP 并向邻居设备泛洪，同时也可以学习到邻居路由设备泛洪而来的 LSP。LSP 同样也分为 Level-1 和 Level-2 两类。由对应级别的 Level-1 与 Level-2 路由器产生，Level-1-2 路由器可以同时产生和接受两种 LSP。SNP 报文通过描述全部或部分数据库中的 LSP 来同步各 LSDB，从而实现相同区域中同级别的 LSDB 的完整与同步。SNP 又可以分为完全序列号 PDU（Complete SNP，CSNP）和部分序列号 PDU（Partial SNP，PSNP）。PSNP 只列举最近收到的一个或多个 LSP 的序号，能一次对多个 LSP 进行确认。同时，设备通过发送 PSNP 来请求更新自己的 LSDB。CSNP 则是包括本地某个级别的 LSDB 中所有的 LSP 信息，可以使相邻路由器保持同级别的 LSDB 同步。

9.3.2　IS-IS 配置

【例 15】 图 9.27 为一个典型的 IS-IS 网络。其中 R0、R1 和 R2 为 Level-2 路由器，他们所属的区域 Area 1 为骨干区域。路由器 R3、R4 为 Level-1-2 路由器，其余的路由器均为 Level-1 路由器，它们所属的区域 Area 2、Area 3 为非骨干区域。可以看见，所有非骨干区域间的通信都要通过 Level-2 区域来进行交换。

答：每个路由器都设置有 isis 地址，如 R0：49.0001 为 area id，0001.0001.0001.00 为 system id。

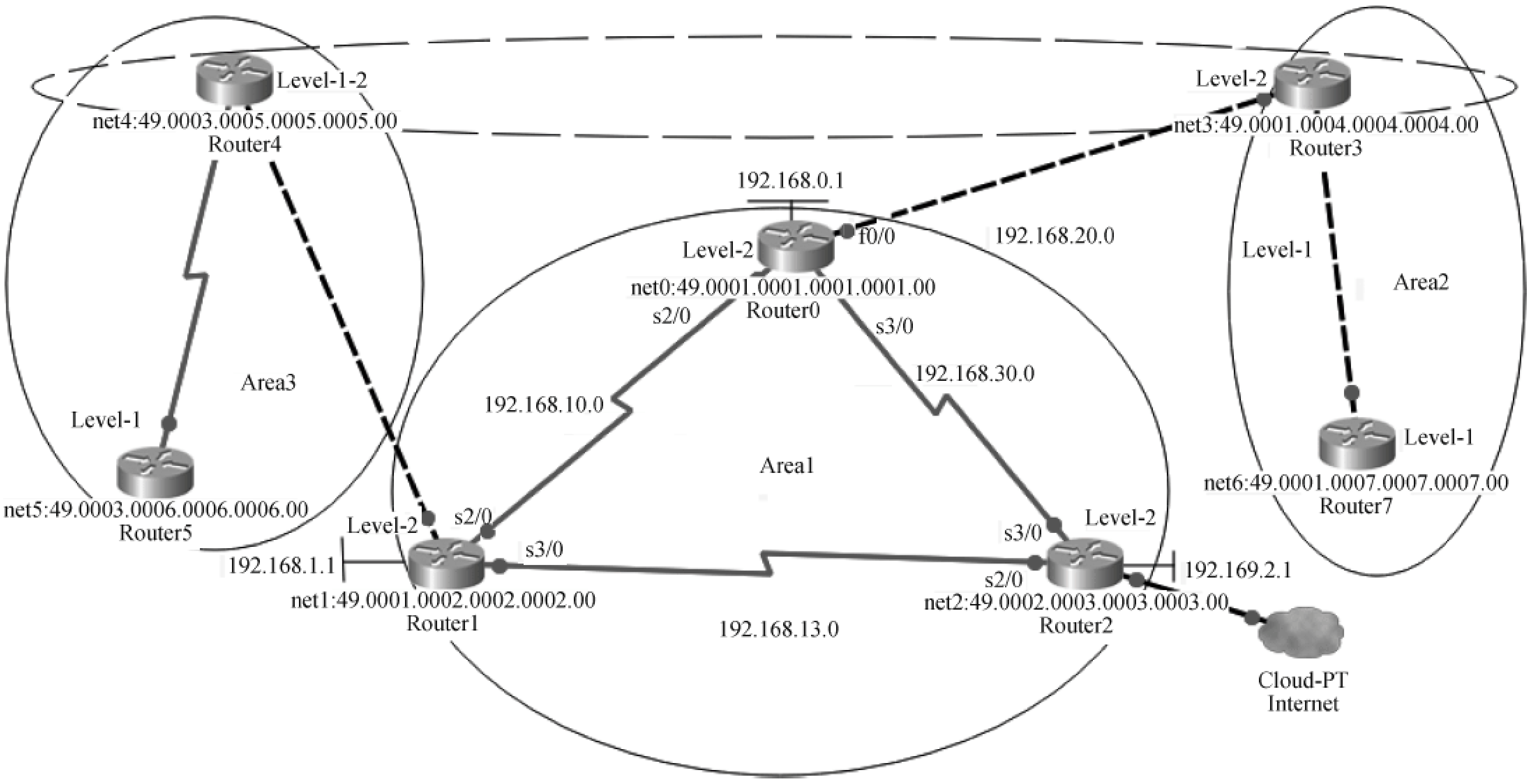


图 9.27　IS-IS 网络拓扑图

Router0 的配置如下：

```
Router(config)#router isis                                   !启用 IS-IS
Router(config-router)net 49.0001.0001.0001.00                !定义 IS-IS 地址
Router (config-router)#isis type level-2                     !配置路由器的类型为 Level-2
Router(config-router)#exit
Router(config)#interface loopback 0
Router(config-if)#ip address 192.168.0.1 255.255.255.0      !配置接口 loopback 0 IP 地址
Router(config-if)ip router is-is              !该接口上激活 IS-IS,使它参与 IS-IS 路由选择
Router(config-if)#isis circuit-type level-2                  !建立 Level-2 级邻居关系
Router(config-if)#no shutdown
Router(config)#interface s3/0
Router(config-if)#ip address 192.168.30.1 255.255.255.0     !配置接口 s3/0 IP 地址
Router(config-if)#clock rate 2000000                         !s3/0 为 DTC 接口
Router(config-if)ip router is-is
Router(config-if)#isis circuit-type level-2                  !s3/0 建立 Level-2 邻居关系
Router(config-if)#no shutdown
Router(config)#interface s2/0
Router(config-if)#ip address 192.168.10.1 255.255.255.0
Router(config-if)#clock rate 2000000                         !s2/0 为 DTC 接口
Router(config-if)ip router is-is
Router(config-if)#isis circuit-type level-2                 !接口 s2/0 建立 Level-2 邻居关系
Router(config-if)#no shutdown
Router(config)#interface f0/0
Router(config-if)#ip address 192.168.20.1 255.255.255.0
Router(config-if)#clock rate 2000000
Router(config-if)ip router is-is
Router(config-if)#isis circuit-type level-2 only          !接口 f0/0 只建立 Level-2 邻居关系
Router(config-if)#no shutdown
```

习题

1. RIP 配置是存放在计算机中,还是路由器中? 它用哪一种传输层协议?
2. 简述 RIP 的工作流程。
3. RIP 都有哪些主要配置命令?
4. 参考图 9.7,练习 RIPv2 配置,并验证配置结果。
5. 简述 OSPF 5 种不同类型的数据包。
6. OSPF 系统是如何进行区域规划的? 并遵守哪些原则?
7. 参考图 9.22,进行多区域 OSPF 配置。
8. 参考图 9.26,进行 STUB 区域、路由聚合和虚连接 OSPF 配置。
9. IS-IS 路由器是如何分类的,说明 Level-1、Level-2 和 Level-1-2 的关系。

第10章 自治区域外部网关协议

目前,边界网关协议是一种应用最广的自治系统间动态路由发现协议,它与自治区域内的 OSPF、RIP 是相互配合的关系。BGP 作为自治区域外部网关协议(EGP),其着眼点在于控制路由的传播和选择最好的路由,而 OSPF、RIP 则主要用于发现和计算路由。本章主要介绍 BGP 路由协议的有关工作原理,以及在自治区域间运行 BGP 的有关配置。

10.1 BGP 工作原理

边界网关协议(Border Gateway Protocol,BGP)发送和引入路由的单位是整个自治系统(Autonomous System,AS),即 BGP 要发送本地路由器所在 AS 内部的所有路由,其路由数量显然要远远大于 IGP 发送和引入的路由数量。因此,类似于 IGP 那样定时对外广播路由信息是不可取的。

10.1.1 BGP 路由

1. BGP 路由选择

一个 AS 就是处于一个管理机构控制之下的路由器和网络群组。指的是由同一个机构管理、使用统一选路策略的一些路由器的集合。BGP 作为 AS 区域间的路由协议,要按照不同的路由的属性控制路由的发送和引入。每个自治系统都有唯一的自治系统编号,自治系统的编号范围是 1～65 535,其中 1～65 411 是注册的因特网编号,65 412～65 535 是专用网络编号。通过采用路由协议和自治系统编号,路由器就可以确定彼此间的路径和路由信息的交换方法。BGP 使用 TCP 作为其传输层协议。

在 BGP 中,拓扑图的端点是一个 AS 区域,端点之间的连接便是链路。IGP 负责在 AS 内部选择花费最小的路由,EGP 负责选择 AS 间花费最小的路由。BGP 采用发送路由增量的方法,完成全部路由信息的通告。

当本地路由器的 BGP 收到了一条新路由时,与保存的已发送信息进行比较,如未发送过,则发送;如已发送过,则与已经发送的路由进行比较,如新路由花费更小,则发送此新路由,同时更新已发送信息,反之则不发送。

当本地路由器 BGP 发现一条路由失效时,如此路由已发送过,则向 BGP 对等体发送一个撤销路由消息。总之,BGP 不是每次都广播所有的路由信息,而是在初始化全部路由信息后只发送路由的变化量(增量),这样保证了 BGP 和对端的最小通信量。因为对于 IGP,本地路由协议只需发送时刻所知的全部路由,而不保存任何已发送信息,路由选择的工作由对端来完成;而 BGP 必须为每个 BGP 对端保存已经发送的路由信息,以便发送一条新路由前确认其是否真的应该发送。

BGP 还支持无类别域间选路(CIDR)。它使用带有较短掩码的路由在一条路由中表达

更多的路由信息，也就是路由聚合。如从 210.11.1.0/24～210.11.254.0/24 可以使用 210.11.0.0/16 表示，从而减小路由表的体积和发送路由的通信量。以下给出路由选择的过程：

如果此路由的下一跳不可达，忽略此路由；

选择本地优先级较大的路由；

选择本地路由器始发的路由(本地优先级相同)；

选择 AS 路径较短的路由；

依次选择起点类型为 IGP、EGP、INCOMPLETE(非完整)类型的路由；

选择多出口区分(Multi-Exit Discriminators，MED)较小的路由；

选择 RouterID 较小的路由。

2. BGP 路由注入

BGP 主要功能是在自治系统之间传递路由信息，它的功能不在于发现和计算路由，BGP 传递的路由信息需要注入。BGP 路由注入方式有 3 种：纯动态注入、半动态注入和静态注入。

纯动态注入：指路由器将通过 IGP 路由协议动态获得的路由信息直接注入 BGP 中去。纯动态注入方式没有对路由信息做任何过滤和选择，它会把路由器获得的所有 IGP 路由信息都引入到 BGP 系统中。从另一角度来说，这样一种路由注入方式配置简单，一次性引入了所有的路由信息。如图 10.1 所示，RouterA 通过 OSPF 协议动态地发现去往网络 10.10.0.0/16 的路由后，就自动引入到 BGP 中。

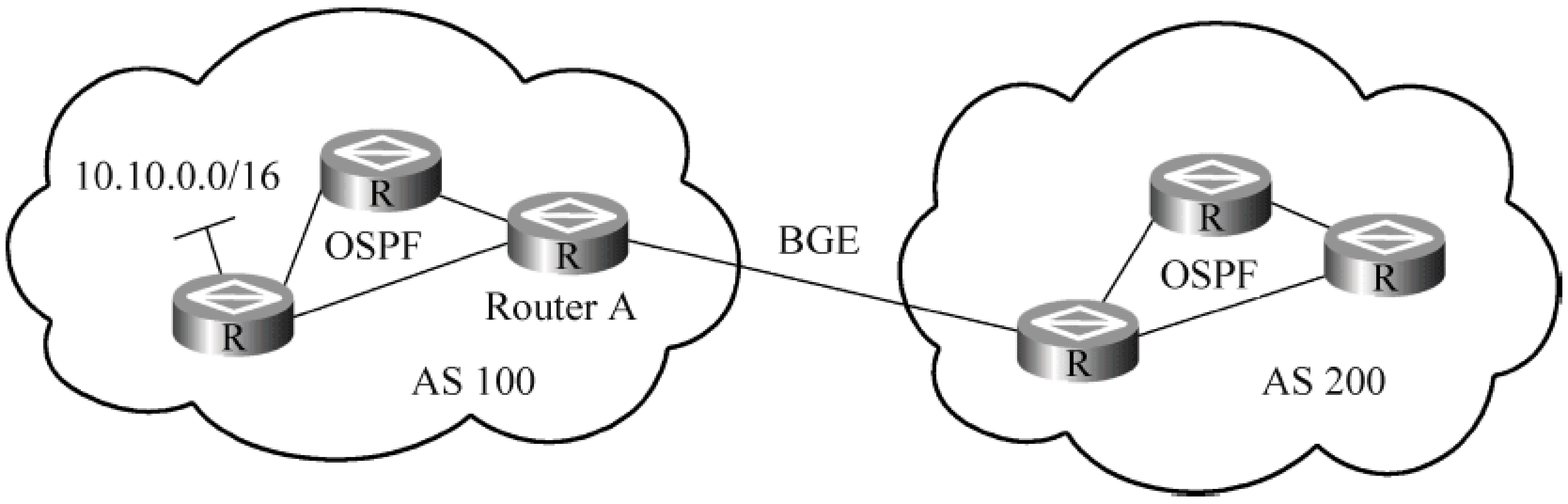


图 10.1 自治系统之间传递路由信息示意图

半动态注入：指路由器有选择性地将 IGP 发现的动态路由信息注入 BGP 系统中去。它和纯动态注入的区别在于不是将 IGP 发现的所有路由信息注入 BGP 中去。如图 10.1 所示，RouterA 通过 OSPF 协议动态地发现去往网络 10.10.0.0/16 的路由，再通过配置命令静态地将其引入到 BGP 中，我们称这样一种路由注入方式为半动态注入。

静态注入：指路由器将静态配置的某条路由注入 BGP 系统中。如图 10.1 所示，RouterA 首先建立一条去往网络 10.10.0.0/16 的静态路由，再通过配置命令将其静态引入到 BGP 中，这样的一种路由注入方式就称为静态注入。

10.1.2 BGP 报文和状态机

1. BGP 报文

1) BGP 报文消息类型

BGP 的运行是通过消息驱动的，BGP 消息使用以下 5 种报文类型。

(1) Type 1,OPEN(开放):TCP 连接建立后发送的第一个消息,用于建立 BGP 对等体间的连接关系。OPEN 消息的语义是:“你好,我们交个朋友好吗?”。BGP 对等体间通过发送 OPEN 报文来交换各自的版本、自治系统号、保持时间、BGP 标识符等信息,相互进行协商。

(2) Type 2,UPDATE(更新):BGP 系统中最重要的信息,用于在对等体之间交换路由信息,它最多由 3 部分构成:不可达路由(unreachable)、路径属性(path attributes)和网络可达性信息(Network Layer Reachability Information,NLRI)。UPDATE 消息的语义是:“有新闻消息通告你”。UPDATE 报文携带的是路由更新信息。其中包括撤销路由信息和可达路由信息及其路径属性等。

(3) Type 3,NOTIFICATION(通知):错误通告消息。NOTIFICATION 消息的语义是:我不跟你玩了。当 BGP 检测到差错(连接中断、协商出错等)时,发送此报文,关闭同对等体的连接。

(4) Type 4,KEEPALIVE(保持激活):用于检测连接有效性的消息。KEEPALIVE 消息的语义是:“我还好着呢,别不理我呀”。KEEPALIVE 报文在 BGP 对等体间周期地发送,以确保连接保持有效。

(5) Type 5,ROUTE-REFRESH(路由更新):用于通知对等体自己支持路由刷新能力。

2) BGP 对等体(peer)

相互交换消息的 BGP 发言人之间互称对等体(peer),若干相关的对等体可以构成对等体组(group)。OPEN 报文是 BGP 路由器之间的初始握手消息,用于建立邻居(BGP 对等体)关系,应该发生在任何通告消息之前。其他路由器在收到 OPEN 消息之后,即以 KEEPALIVE 消息作为响应。BGP 初次启动时,路由器发送整个 BGP 路由表与对等体交换路由信息,之后只交换 UPDATE(更新消息)。运行过程中,通过接收和发送 KEEPALIVE(保持激活)消息检测相互之间的连接是否正常,以及 NOTIFICATION 等消息的交换操作。

发送 BGP 消息的路由器称为 BGP 发言人(speaker),它接收或产生新的路由信息,并发布给其他 BGP 发言人。当 BGP 发言人收到来自其他自治系统的新路由时,如果该路由比当前已知路由更优,或者当前还没有该路由,它就把这条路由发布给自治系统内所有其他 BGP 发言人。

3) BGP 消息的具体应用过程

(1) BGP 使用 TCP 建立连接,本地监听端口号为 179。和 TCP 建立相同,BGP 连接的建立也要经过一系列的对话和握手。BGP 的握手协商的参数有:BGP 版本、BGP 连接保持时间、本地的路由器标识(Router ID)、授权信息等。这些信息都在 Open 消息中体现。

(2) BGP 连接建立后,如果有路由需要发送则发送 UPDATE 消息通告对端路由信息。UPDATE 消息主要用来通告路由信息,包括失效(撤销)路由。UPDATE 消息发布路由时,还要指定此路由的路由属性,用以帮助对端 BGP 选择最佳的路由。

(3) 在本地 BGP 路由变化时,也使用 UPDATE 消息修正对端 BGP 的路由表。

(4) 经过一段时间的路由信息交换后,本地 BGP 和对端 BGP 都无新路由通告,就是趋于稳定了。此时要定时发送 KEEPALIVE 消息以保持 BGP 连接的有效性。对于本地 BGP,如果在超过保持时间还未收到任何对端 BGP 消息,就认为此 BGP 连接已经无效,将断开。

(5) 当本地 BGP 在运行中发现错误时,要发送 NOTIFY 消息通告 BGP 对端。如对端 BGP 版本本地不支持,本地 BGP 收到了结构非法的 UPDATE 消息等。本地 BGP 退出 BGP 连接时也要发送 NOTIFICATION 消息。BGP 收到 NOTIFICATION 消息后,要做相应处理。

2. BGP 的状态机

BGP 有限状态机有 6 个状态,BGP 的状态机如图 10.2 所示,它们之间的转换过程示意了 BGP 邻居关系建立的过程。以下给出这 6 个状态。

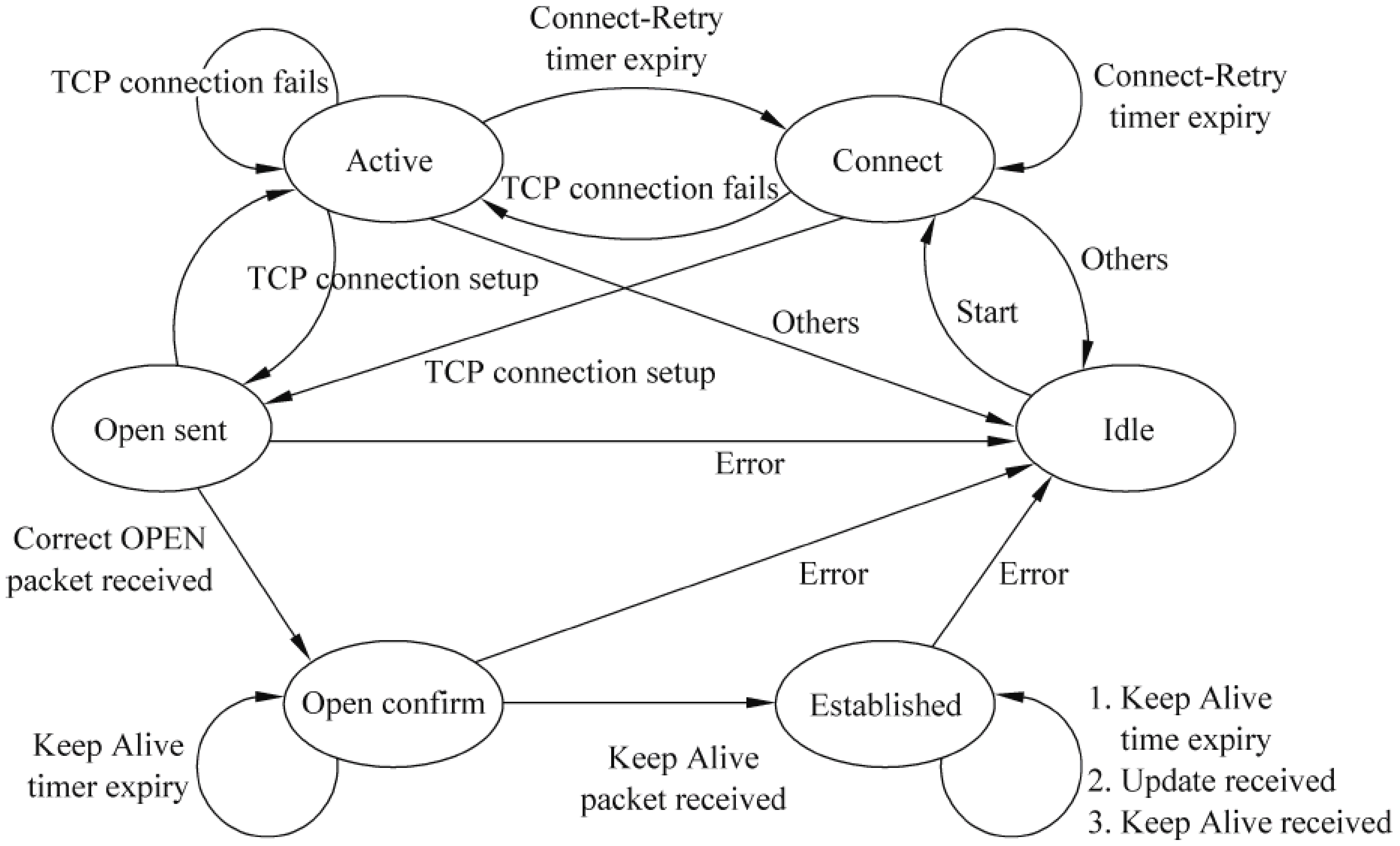


图 10.2 BGP 的状态机

(1) Idle(空闲):BGP 连接的第一个状态,在空闲状态,BGP 在等待一个启动事件,启动事件出现以后,BGP 初始化资源,复位连接重试计时器(Connect-Retry),发起一条 TCP 连接,同时转入 Connect 状态。

(2) Connect(连接):在 Connect 状态,BGP 发起第一个 TCP 连接,如果连接重试计时器(Connect-Retry)超时,就重新发起 TCP 连接,并继续保持在 Connect 状态,如果 TCP 连接成功,就转入 Opensent 状态;如果 TCP 连接失败,就转入 Active 状态。

(3) Active(激活):在 Active 状态,BGP 总是在试图建立 TCP 连接,如果连接重试计时器(Connect-Retry)超时,就退回到 Connect 状态,如果 TCP 连接成功,就转入 Opensent 状态,如果 TCP 连接失败,就继续保持在 Active 状态,并继续发起 TCP 连接。

(4) Opensent(打开消息已发送):在 Opensent 状态,TCP 连接已经建立,BGP 也已经发送了第一个 Open 报文,剩下的工作,BGP 就在等待其对等体发送 Open 报文,并对收到的 Open 报文进行正确性检查,如果有错误,系统就会发送一条出错通知消息并退回到 Idle 状态;如果没有错误,BGP 就开始发送 Keepalive 报文,并复位 Keepalive 计时器,开始计时。同时转入 Openconfirm 状态。

(5) Openconfirm(打开消息确认)状态:在 Openconfirm 状态,BGP 等待一个 Keepalive 报文,同时复位保持计时器,如果收到了一个 Keepalive 报文,就转入 Established 阶段,BGP 邻居关系就建立起来了。如果 TCP 连接中断,就退回到 Idle 状态。

(6) Established(连接已建立):在 Established 状态,BGP 邻居关系已经建立,这时,BGP 将和它的邻居们交换 Update 报文,同时复位保持计时器。

综上状态机的转换过程首先是在 Idle 状态,BGP 一旦 Start,状态机就进入 Connect 状态,在 Connect 状态,如果 Connect-Retry 定时器超时,BGP 状态机会停留在 Connect 状态,如果 TCP 连接建立成功,BGP 状态机就直接进入 Opensent 状态。在 Opensent 状态,BGP 一旦收到了一个正确的 Open 报文,就会进入 Openconfirm 状态。在 Openconfirm 状态,如果 Keepalive 定时器超时,BGP 状态机就会停留在 Openconfirm 状态。直到 BGP 收到 Keepalive 报文,BGP 状态机才会进入 Established 状态。这时 BGP 连接才算建立起来。另外,在除 Idle 状态以外的其他 5 个状态出现任何 Error 的时候,BGP 状态机就会退回到 Idle 状态。

10.2 BGP 配置

10.2.1 BGP 基本配置

以下首先给出华为(Quidway)系列路由器主要的配置命令和过程。

1. BGP 配置命令

1) 启动 BGP

```
bgp as - number
```

as-number 是指本地的自治系统号,路由器在同一时间只允许启动一个 BGP 进程。因此,一台路由器只能属于一个自治系统。undo bgp 命令用来关闭 BGP。

2) 配置 IBGP 邻居

启动了 BGP 进程,同时进入 BGP 配置模式。下一步就是配置 BGP 邻居,IBGP (Internal BGP)和 EBGP(External BGP)就是 BGP 的两种邻居。当 BGP 运行于同一自治系统内部时,称为 IBGP;当 BGP 运行于不同自治系统之间时,称为 EBGP。配置 BGP 邻居命令如下。

```
peer {group - name | ipv4 - address | ipv6 - address}as - number as - number
```

其中,group-name:对等体组的名称;

ipv4-address:对等体的 IPv4 地址;

ipv6-address:对等体的 IPv6 地址;

as-number:对等体/对等体组的对端路由器 AS 号,如果需要配置 EBGP 邻居,as-number 就是本端路由器所在的自治系统号。

相互交换消息的 BGP 邻居之间互称对等体(peer),若干相关的对等体可以构成对等体组(group)。如:执行 peer as-number 命令配置指定对等体(组)的对端 AS 号;执行 undo peer as-number 命令来删除对等体(组)的 AS 号。默认情况下,对等体组对端无 AS 号。

例如,配置 IPv4 对等体 2.2.2.2 的对端 AS 号为 8:

```
<Quidway> system - view
[Quidway] bgp 8
```

```
[Quidway-bgp] peer 2.2.2.2 as-number 8
```

例如,配置 IPv6 对等体 2::2 的对端 AS 号为 8:

```
[Quidway-bgp] peer 2::2 as-number 8
```

3) BGP 注入路由

随着 BGP 进程的启动与邻居关系的建立,BGP 路由协议也就运行起来了。但 BGP 路由表中没有任何路由信息,因此要为 BGP 系统注入路由信息可用命令:

```
network ip-address [ mask mask ]
```

或

```
import-route protocol
```

protocol 指定可引入的源路由协议,可以是 direct、static、rip、ospf 等。

2. BGP 配置主要过程

如图 10.3 所示,AS20 内每两台路由器之间建立 BGP 邻居关系,RTB、RTC 分别和 RTA、RTD 建立 EBGP 邻居关系,相互通告路由信息。在 AS20 内部运行 OSPF 路由协议。

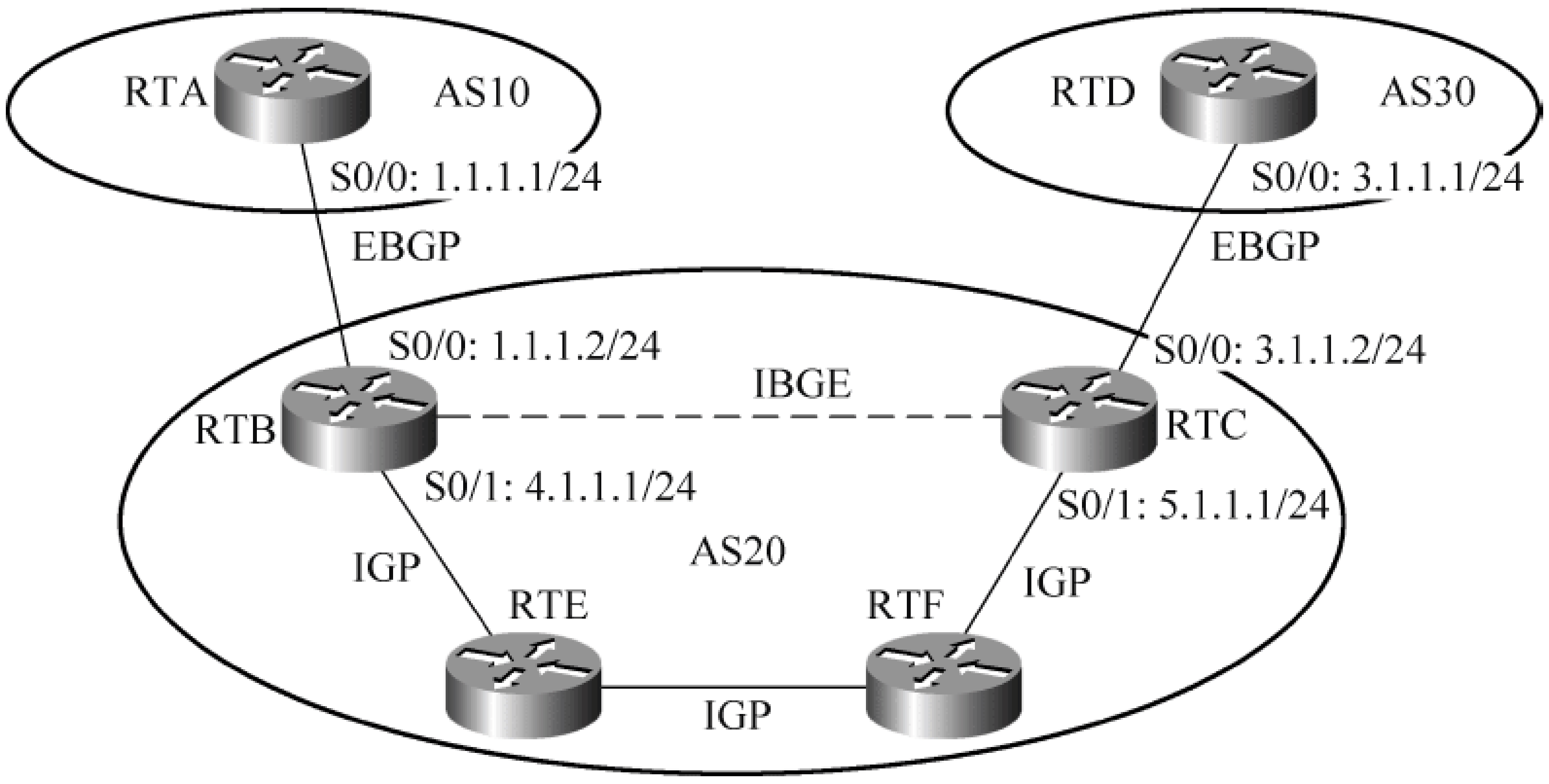


图 10.3 BGP 基本配置示意图

已知 BGP 在路由器上有两种运行方式:IBGP 和 EBGP。如果两个交换 BGP 报文的对等体属于同一个自治系统,那么这两个对等体就是 IBGP 对等体,例如,RTB 和 RTC。

如果两个交换 BGP 报文的对等体属于不同的自治系统,那么这两个对等体就是 EBGP 对等体(External BGP),如 RTA 和 RTB,或 RTD 和 RTC。

虽然 BGP 是运行于自治系统之间的路由协议,但是一个 AS 的不同边界路由器之间也要建立 BGP 连接,只有这样才能实现路由信息在全网的传递,如 RTB 和 RTC。为了建立 AS10 和 AS30 之间的通信,就需要在它们之间建立 IBGP 连接。

IBGP 对等体之间不一定是物理上直连的,但必须保证逻辑上全连接(通过 TCP 连接能够建立即可)。一般情况下,路由器都默认要求 EBGP 对等体之间有物理上的直连链路,同时它们一般也提供改变这个默认设置的配置命令。

如前所述,BGP 基本配置包括 3 步:首先启动 BGP,然后配置 BGP 邻居,最后为 BGP 注入路由信息。以下根据图 10.3 所示,列出 Quidway 系列路由器的主要配置过程。

(1) 配置 RTA。

```
[RTA- Serial0/0]ip address 1.1.1.1 255.255.255.0   !为接口 S0/0 配置 IP 地址
[RTA]bgp 10                                         !启动 BGP,AS 为 10
[RTA-bgp]peer 1.1.1.2 as-number 20                  !与 RTB 建立 EBGP 邻居
[RTA-bgp]import-route direct                        !导入直连路由,为 BGP 注入路由信息
```

(2) 配置 RTB。

```
[RTB- Serial0/0]ip address 1.1.1.2 255.255.255.0   !配置 S0/0 接口 IP 地址
[RTB- Serial0/1]ip address 4.1.1.1 255.255.255.0   !配置 S0/1 接口 IP 地址
[RTB] ospf enable                                   !启动 ospf
[RTB- Serial0/0]ospf enable area 0                  !配置 S0/0 属于 OSPF area0
[RTB- Serial0/1]ospf enable area 0                  !配置 S0/1 属于 OSPF area0
[RTB] bgp 20                                        !启动 BGP,AS 为 20
[RTB-bgp]peer 1.1.1.1 as-number 10                  !与 RTA 建立 EBGP 邻居
[RTB-bgp]peer 5.1.1.1 as-number 20                  !与 RTC 建立 IBGP 邻居
[RTB-bgp] import-route ospf                         !导入动态路由协议 OSPF,为 BGP 注入路由信息
```

(3) 配置 RTC。

```
[RTC-Serial0/0]ip address 3.1.1.2 255.255.255.0    !配置 S0/0 接口 IP 地址
[RTC-Serial0/1]ip address 5.1.1.1 255.255.255.0    !配置 S0/1 接口 IP 地址
[RTC]ospf enable                                    !启动 ospf
[RTC-Serial0/0]ospf enable area 0                   !配置 S0/0 属于 OSPF area0
[RTC-Serial0/1]ospf enable area 0                   !配置 S0/1 属于 OSPF area0
[RTC] bgp 20                                        !启动 BGP,AS 为 20
[RTC-bgp]peer 3.1.1.1 as-number 30                  !与 RTD 建立 EBGP 邻居
[RTC-bgp]peer 4.1.1.1 as-number 20                  !与 RTB 建立 IBGP 邻居
[RTC-bgp]import-route ospf                          !导入动态路由协议 OSPF,为 BGP 注入路由信息
```

(4) 配置 RTD。

```
[RTD-Serial0/0]ip address 3.1.1.1 255.255.255.0    !配置 S0/0 接口 IP 地址
[RTD]bgp 30                                         !启动 BGP,AS 为 30
[RTD-bgp]peer 3.1.1.1 as-number 20                  !与 RTC 建立 EBGP 邻居
[RTD-bgp] import-route direct                       !导入直连路由,为 BGP 注入路由信息
```

10.2.2 BGP 配置实例

下面要熟悉与 Quidway 命令功能相似的 3 个 Cisco 配置命令,在实例中要用到。

```
router bgp as-num                           !启动 BGP,as-num 为 AS 号
network ip-address mask mask                !指定与路由器相连的网段,运行 BGP
neighbor ip-address remote-as as-num        !指定与对等体路由器相连的接口地址,并与其所
                                            !在的 AS 建立 IBGP 或 EBGP 邻居
```

在以上命令中,ip-address 为 IP 地址;mask 为掩码;as-num 为路由器 AS 号。

1. BGP 详细配置

BGP 配置拓扑结构如图 10.4 所示,共分为 3 个 AS,即 AS100、AS200 和 AS300。在各

台路由器上配置基本的 IP 地址，包括回环地址。在路由器 B、C、D 上要配置 RIP2 协议，相互之间通告 192.168.0.0 网络路由信息，并在有关路由器上启动 BGP。以下给出路由器 B 的详细配置及注释，其他路由器的详细配置只给出参考截图。

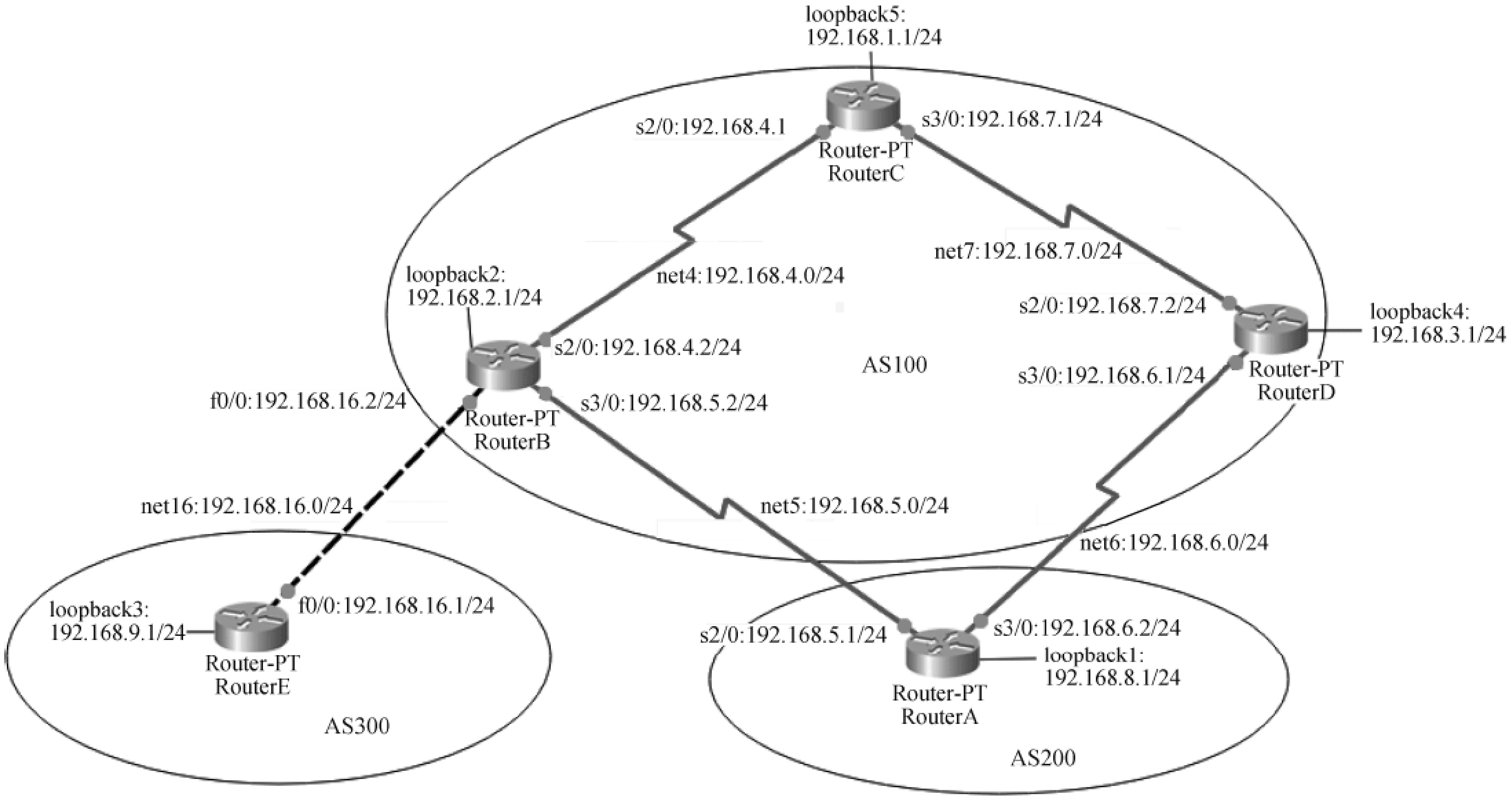


图 10.4　BGP 配置拓扑结构

(1) 首先在各台路由器上通过 hostname 修改名字，另外，为了在故障排除(troubleshooting)时便于检测，根据实际需求，可以配置虚拟终端。

```
Router> enable
Router# configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Router(config)# hostname AS100-B                !路由器 RouterB 配置 hostname 为 AS100-B
AS100-B(config)# line vty 0 4
AS100-B(config-line)# password cisco123         !设置路由器使能密码
AS100-B(config-line)# exec-timeout 4 30         !设置 4 分 30 秒内没有键盘输入，则强行断开
AS100-B(config-line)# login                     !设置在登录终端线时，需要检查密码
AS100-B(config-line)# exit
AS100-B (config)
```

(2) 在各路由器上配置基本的 IP 地址，包括回环地址，并配置 DCE 时钟。

```
AS100-B(config)# interface loopback2            !设定回环接口 Loopback2 的 IP 地址
AS100-B(config-if)#
%LINK-5-CHANGED: Interface Loopback2, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface Loopback2, changed state to up
AS100-B(config-if)# ip address 192.168.2.1 255.255.255.0
AS100-B(config-if)# exit
AS100-B(config)# interface f0/0                 !配置 f0/0 的 IP 地址
AS100-B(config-if)# ip address 192.168.16.2 255.255.255.0
AS100-B(config-if)# no shutdown                 !激活 f0/0 接口
AS100-B(config-if)#
%LINK-5-CHANGED: Interface FastEthernet0/0, changed state to up
AS100-B(config-if)# exit
```

```
AS100-B(config)#interface s2/0                    !配置 s2/0 的 IP 地址
AS100-B(config-if)#ip address 192.168.4.2 255.255.255.0
AS100-B(config-if)#clock rate 2000000             !设置 DCE 接口时钟频率
AS100-B(config-if)#no shutdown                    !激活 s0/0 接口
%LINK-5-CHANGED: Interface Serial2/0, changed state to down
AS100-B(config-if)#exit
AS100-B(config)#interface s3/0                    !配置 s3/0 的 IP 地址
AS100-B(config-if)#ip address 192.168.5.2 255.255.255.0
AS100-B(config-if)#clock rate 2000000             !设置 DCE 接口时钟频率
AS100-B(config-if)#no shutdown                    !激活 s3/0 接口
%LINK-5-CHANGED: Interface Serial3/0, changed state to down
AS100-B(config-if)#exit
```

(3) 在各路由器上要配置 RIP version2 协议,相互之间通告网络路由信息。

```
AS100-B(config)#route rip                           !配置 RIP version2 协议
AS100-B(config-router)#version 2
AS100-B(config-router)#network 192.168.4.0          !指定路由器 RouterB、RouterC 运行网络
AS100-B(config-router)#network 192.168.7.0          !指定路由器 RouterD、RouterC 运行网络
AS100-B(config-router)#network 192.168.3.0          !指定路由器 RouterD 的 Loopback4
AS100-B(config-router)#network 192.168.2.0          !指定路由器 RouterB 的 Loopback2
AS100-B(config-router)#network 192.168.1.0          !指定路由器 RouterC 的 Loopback5
AS100-B(config-router)#exit          !以上配置可以用一条汇聚路由 network 192.168.0.0 代替
```

(4) 在 RouterA、RouterB、RouterD、RouterE 边缘路由器上都启动 BGP 并进行相关配置。

```
AS100-B(config)#router bgp 100          !配置 BGP,并在与 RouterB 的相连的网段上启动 BGP
AS100-B(config-router)#network 192.168.16.0 mask 255.255.255.0
                                                    !指定 RouterB 与 RouterE 直连网段
AS100-B(config-router)#network 192.168.5.0 mask 255.255.255.0
                                                    !指定 RouterB 与 RouterA 直连网段
AS100-B(config-router)#neighbor 192.168.5.1 remote-as 200
                                                    !RouterB 与 RouterA 建立 EBGP 邻居
AS100-B(config-router)#
%BGP-3-NOTIFICATION: received from neighbor 192.168.5.1 2/2 (peer in wrong AS) 2 bytes 0064
AS100-B(config-router)#neighbor 192.168.16.1 remote-as 300
                                                    !RouterB 与 RouterE 建立 EBGP 邻居
AS100-B(config-router)# %BGP-5-ADJCHANGE: neighbor 192.168.16.1 Up
AS100-B(config-router)#
%BGP-3-NOTIFICATION: received from neighbor 192.168.5.1 2/2 (peer in wrong AS) 2 bytes 0064
AS100-B(config-router)#exit
AS100-B(config)#exit
```

(5) RouterA 配置截屏如图 10.5 所示,RouterC 配置截屏如图 10.6 所示,RouterD 配置截屏如图 10.7 所示,RouterE 配置截屏如图 10.8 所示。

2. 查看 BGP 配置

根据图 10.4 完成各台路由器配置后,用 show run 等命令看配置正确与否以及相关信息。

```
AS100-B#show run                                    !查看配置是否正确
AS100-B#show ip route                               !查看相应的路由信息
AS100-B#show ip bgp                                 !查看 BGP 信息
AS100-B#show ip bgp neighbors                       !查看 BGP 相邻信息
```

```
AS200-A#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
AS200-A(config)#interface loopback 1
AS200-A(config-if)#ip address  192.168.8.1 255.255.255.0
AS200-A(config-if)#exit
AS200-A(config)#interface s2/0
AS200-A(config-if)#ip address 192.168.5.1 255.255.255.0
AS200-A(config-if)#no shutdow
AS200-A(config-if)#exit
AS200-A(config)#interface s3/0
AS200-A(config-if)#ip address 192.168.6.2 255.255.255.0
AS200-A(config-if)#no shutdown
AS200-A(config-if)#exit
AS200-A(config)#router rip
AS200-A(config-router)#version 2
AS200-A(config-router)#network 192.168.8.0
AS200-A(config-router)#network 192.168.5.0
AS200-A(config-router)#network 192.168.6.0
AS200-A(config-router)#exit
AS200-A(config)#router bgp 200
AS200-A(config-router)#network 192.168.5.0 mask 255.255.255.0
AS200-A(config-router)#network 192.168.6.0 mask 255.255.255.0
AS200-A(config-router)#neighbor 192.168.5.2 remote-as 100
AS200-A(config-router)#%BGP-5-ADJCHANGE: neighbor 192.168.5.2 Up
AS200-A(config-router)#neighbor 192.168.6.1 remote-as 100
AS200-A(config-router)#%BGP-5-ADJCHANGE: neighbor 192.168.6.1 Up
AS200-A(config-router)#exit
AS200-A(config)#
```

图 10.5　RouterA 配置截屏

```
Router>
Router>enable
Router#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
Router(config)#hostname AS100-C
AS100-C(config)#interface loopback5
AS100-C(config-if)#
%LINK-5-CHANGED: Interface Loopback5, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface Loopback5, changed state
to up
AS100-C(config-if)#ip address 192.168.1.1 255.255.255.0
AS100-C(config-if)#exit
AS100-C(config)#interface s2/0
AS100-C(config-if)#ip address 192.168.4.1 255.255.255.0
AS100-C(config-if)#no shutdown
AS100-C(config-if)#
%LINK-5-CHANGED: Interface Serial2/0, changed state to up
AS100-C(config-if)#
%LINEPROTO-5-UPDOWN: Line protocol on Interface Serial2/0, changed state
to up
AS100-C(config-if)#exit
AS100-C(config)#interface s3/0
AS100-C(config-if)#ip address 192.168.7.1 255.255.255.0
AS100-C(config-if)#clock rate 2000000
AS100-C(config-if)#no shutdown
%LINK-5-CHANGED: Interface Serial3/0, changed state to down
AS100-C(config-if)#
AS100-C con0 is now available
Press RETURN to get started.
AS100-C>enable
AS100-C#configur terminal
Enter configuration commands, one per line.  End with CNTL/Z.
AS100-C(config)#router rip
AS100-C(config-router)#version 2
AS100-C(config-router)#network 192.168.3.0
AS100-C(config-router)#network 192.168.2.0
AS100-C(config-router)#network 192.168.4.0
AS100-C(config-router)#network 192.168.1.0
AS100-C(config-router)#network 192.168.7.0
AS100-C(config-router)#exit
AS100-C(config)#ip route 0.0.0.0 0.0.0.0 192.168.4.2
AS100-C(config)#
```

图 10.6　RouterC 配置截屏

```
Router>enable
Router#configure terminal
Enter configuration commands, one per line.  End with CNTL/Z.
Router(config)#hostname AS100-D
AS100-D(config)#interface loopback4

AS100-D(config-if)#
%LINK-5-CHANGED: Interface Loopback4, changed state to up

%LINEPROTO-5-UPDOWN: Line protocol on Interface Loopback4, changed state
to up
AS100-D(config-if)#ip address 192.168.3.1 255.255.255.0
AS100-D(config-if)#exit
AS100-D(config)#interface s2/0
AS100-D(config-if)#ip address 192.168.7.2 255.255.255.0
AS100-D(config-if)#no shutdown

AS100-D(config-if)#
%LINK-5-CHANGED: Interface Serial2/0, changed state to up
AS100-D(config-if)#
%LINEPROTO-5-UPDOWN: Line protocol on Interface Serial2/0, changed state
to up
AS100-D(config-if)#exit
AS100-D(config)#interface s3/0
AS100-D(config-if)#ip address 192.168.6.1 255.255.255.0
AS100-D(config-if)#clock rate 2000000
AS100-D(config-if)#no shutdown
%LINK-5-CHANGED: Interface Serial3/0, changed state to down
AS100-D(config-if)#exit
AS100-D(config)#router rip
AS100-D(config-router)#version 2
AS100-D(config-router)#network 192.168.3.0
AS100-D(config-router)#network 192.168.2.0
AS100-D(config-router)#network 192.168.4.0
AS100-D(config-router)#network 192.168.6.0
AS100-D(config-router)#network 192.168.1.0
AS100-D(config-router)#network 192.168.7.0
AS100-D(config-router)#exit
AS100-D(config)#router bgp 100
AS100-D(config-router)#network 192.168.7.0 mask 255.255.255.0
AS100-D(config-router)#neighbor 192.168.6.2 remote-as 200
AS100-D(config-router)#neighbor 192.168.4.2 remote-as 300
AS100-D(config-router)#ip route 0.0.0.0 0.0.0.0 192.168.7.1
AS100-D(config-router)#exit
AS100-D(config)#
```

图 10.7　RouterD 配置截屏

```
AS300-E(config)#interface loopback 3
AS300-E(config-if)#ip address 192.168.9.1 255.255.255.0
AS300-E(config-if)#exit
AS300-E(config)#interface f0/0
AS300-E(config-if)#ip address 192.168.16.1 255.255.255.0
AS300-E(config-if)#no shutdown
AS300-E(config-if)#exit
AS300-E(config)#router rip
AS300-E(config-router)#version 2
AS300-E(config-router)#network 192.168.9.0
AS300-E(config-router)#network 192.168.16.0
AS300-E(config-router)#exit
AS300-E(config)#router bgp 300
AS300-E(config-router)#network 192.168.16.0 mask 255.255.255.0
AS300-E(config-router)#neighbor 192.168.16.2 remote-as 100
AS300-E(config-router)#%BGP-5-ADJCHANGE: neighbor 192.168.16.2 Up
AS300-E(config-router)#neighbor 192.168.6.1 remote-as 100
AS300-E(config-router)#%BGP-5-ADJCHANGE: neighbor 192.168.6.1 Up
AS300-E(config-router)#exit
AS300-E(config)#ip route 0.0.0.0 0.0.0.0 192.168.16.1
%Invalid next hop address (it's this router)
AS300-E(config)#exit
```

图 10.8　RouterE 配置截屏

图 10.9 就是用 show run 命令看到的 RouterB 的有关配置信息。

```
AS100-B#show run
Building configuration...
interface Loopback2
 ip address 192.168.2.1 255.255.255.0
!
interface FastEthernet0/0
 ip address 192.168.16.2 255.255.255.0
 duplex auto
 speed auto
!
interface Serial2/0
 ip address 192.168.4.2 255.255.255.0
 clock rate 2000000
!
interface Serial3/0
 ip address 192.168.5.2 255.255.255.0
 clock rate 2000000
!
router bgp 100
 bgp log-neighbor-changes
 no synchronization
 neighbor 192.168.5.1 remote-as 200
 neighbor 192.168.16.1 remote-as 300
 network 192.168.16.0
 network 192.168.5.0
!
router rip
 version 2
 network 192.168.1.0
 network 192.168.2.0
 network 192.168.3.0
 network 192.168.4.0
 network 192.168.7.0
!
line con 0
!
line aux 0
!
line vty 0 4
 exec-timeout 4 30
 password cisco123
 login
```

图 10.9　查看 RouterB 有关运行结果部分截屏

图 10.10 给出的是用 show ip route 命令查看 RouterA～RouterE 的路由表信息。

```
AS100-B>
AS100-B>enable
AS100-B#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS
inter area
       * - candidate default, U - per-user static route, o - ODR
       P - periodic downloaded static route

Gateway of last resort is not set

R    192.168.1.0/24 [120/1] via 192.168.4.1, 00:00:06, Serial2/0
C    192.168.2.0/24 is directly connected, Loopback2
R    192.168.3.0/24 [120/2] via 192.168.4.1, 00:00:06, Serial2/0
C    192.168.4.0/24 is directly connected, Serial2/0
C    192.168.5.0/24 is directly connected, Serial3/0
B    192.168.6.0/24 [20/0] via 192.168.5.1, 00:48:54
R    192.168.7.0/24 [120/1] via 192.168.4.1, 00:00:06, Serial2/0
R    192.168.8.0/24 [120/3] via 192.168.4.1, 00:00:06, Serial2/0
B    192.168.9.0/24 [20/0] via 192.168.16.1, 00:48:54
C    192.168.16.0/24 is directly connected, FastEthernet0/0
AS100-B#
```

(a) RouterB路由表

```
AS200-A#show ip route
Gateway of last resort is 192.168.5.2 to network 0.0.0.0

R    192.168.1.0/24 [120/2] via 192.168.6.1, 00:00:02, Serial3/0
R    192.168.2.0/24 [120/3] via 192.168.6.1, 00:00:02, Serial3/0
R    192.168.3.0/24 [120/1] via 192.168.6.1, 00:00:02, Serial3/0
R    192.168.4.0/24 [120/2] via 192.168.6.1, 00:00:02, Serial3/0
C    192.168.5.0/24 is directly connected, Serial2/0
C    192.168.6.0/24 is directly connected, Serial3/0
B    192.168.7.0/24 [20/0] via 192.168.6.1, 00:34:45
C    192.168.8.0/24 is directly connected, Loopback1
B    192.168.9.0/24 [20/0] via 192.168.5.2, 00:34:45
B    192.168.16.0/24 [20/0] via 192.168.5.2, 00:34:45
S*   0.0.0.0/0 [1/0] via 192.168.5.2
AS200-A#
```

(b) RouterA路由表

```
AS100-C#show ip route
C    192.168.1.0/24 is directly connected, Loopback5
R    192.168.2.0/24 [120/1] via 192.168.4.2, 00:00:23, Serial2/0
R    192.168.3.0/24 [120/1] via 192.168.7.2, 00:00:25, Serial3/0
C    192.168.4.0/24 is directly connected, Serial2/0
R    192.168.5.0/24 [120/2] via 192.168.7.2, 00:00:25, Serial3/0
R    192.168.6.0/24 [120/1] via 192.168.7.2, 00:00:25, Serial3/0
C    192.168.7.0/24 is directly connected, Serial3/0
R    192.168.8.0/24 [120/2] via 192.168.7.2, 00:00:25, Serial3/0
S*   0.0.0.0/0 [1/0] via 192.168.4.2
AS100-C#
```

(c) RouterC路由表

```
AS100-D#show ip route
R    192.168.1.0/24 [120/1] via 192.168.7.1, 00:00:13, Serial2/0
R    192.168.2.0/24 [120/2] via 192.168.7.1, 00:00:13, Serial2/0
C    192.168.3.0/24 is directly connected, Loopback4
R    192.168.4.0/24 [120/1] via 192.168.7.1, 00:00:13, Serial2/0
B    192.168.5.0/24 [20/0] via 192.168.6.2, 00:48:09
C    192.168.6.0/24 is directly connected, Serial3/0
C    192.168.7.0/24 is directly connected, Serial2/0
R    192.168.8.0/24 [120/1] via 192.168.6.2, 00:00:15, Serial3/0
B    192.168.9.0/24 [20/0] via 192.168.16.1, 00:48:09
B    192.168.16.0/24 [20/0] via 192.168.16.1, 00:48:09
S*   0.0.0.0/0 [1/0] via 192.168.7.1
AS100-D#
```

(d) RouterD路由表

```
AS300-E#show ip route
Gateway of last resort is 192.168.16.2 to network 0.0.0.0

B    192.168.5.0/24 [20/0] via 192.168.16.2, 00:41:18
B    192.168.6.0/24 [20/0] via 192.168.16.2, 00:41:18
B    192.168.7.0/24 [20/0] via 192.168.6.1, 00:41:18
C    192.168.9.0/24 is directly connected, Loopback3
C    192.168.16.0/24 is directly connected, FastEthernet0/0
S*   0.0.0.0/0 [1/0] via 192.168.16.2
AS300-E#
```

(e) RouterE路由表

图 10.10　查看 RouterA～RouterE 的路由表部分截屏

习题

1. 举例说明 BGP 的两种邻居关系：IBGP 和 EBGP。
2. 说明 BGP 的状态机 6 个转换过程。
3. 依据图 10.10，查看 RouterB 的路由信息，要求写出主要意思。
4. 实验题，依据图 10.4，在网络实验室或模拟器上完成配置，并测试运行结果。

第 11 章 NOS与服务器设置

网络操作系统(Network Operating System,NOS),是计算机网络的神经系统,它的基本任务是用统一的方法管理各主机之间的通信和共享资源的利用。本章将围绕 NOS,介绍 C/S(客户/服务)模式、编程接口(API)、套接字(Socket)接口,以及各种 Windows Server 的设置。

11.1 网络操作系统

11.1.1 操作系统简介

1. 操作系统组成及功能

1) 网络操作系统组成

网络操作系统软件可分为 4 部分。

(1) 网络环境软件:配置于服务器上,它使高速并发执行的多任务具有良好的网络环境;它管理工作站与服务器之间的数据传送;提供高速的多用户文件系统。

(2) 网络管理软件:是用于网络管理的操作软件,分为安全性管理软件、容错管理软件、备份软件和性能监测软件。

(3) 工作站网络软件:配置于工作站上,它能实现客户与服务器的交互,使工作站上的用户能访问文件服务器的文件、共享资源。工作站网络软件主要有重定向程序和网络基本输入输出系统。

(4) 网络服务软件:配置在系统服务器上或工作站上,是面向用户的。操作系统是否受到用户的欢迎,主要取决于 NOS 所提供的网络服务软件是否完善。

常用的网络服务软件主要有以下几种。

① 多用户文件服务软件:它为用户程序对服务器中的目录和文件进行有效访问提供了手段,即先由用户向服务器提出文件服务请求,然后由工作站网络服务软件将该请求传送给服务器。该软件既能让多用户共享目录和文件,又限制两个以上工作站不能同时访问某一存储空间,保证数据的安全性。

② 名字服务软件:用户管理网络上所有对象的名字,如进程名、服务器名、各种资源名、文件和目录名等。当用户要访问某一对象时,只需给出该对象的名字即可,并不需要知道该对象的物理地址,名字服务软件能实现寻址和定位服务。

③ 打印服务软件:是将用户的打印信息在服务器上产生假脱机文件,并送到打印机队列中等待打印的软件。

④ 电子邮件服务软件:工作站用户利用该软件把邮件发送给网中其他工作站的用户,实现多址、多地、广播式电子邮件服务。

2）网络操作系统功能

网络操作系统主要功能如下。

(1) 网络通信：网络通信的主要任务是提供通信双方之间无差错的、透明的数据传输服务。主要功能包括建立和拆除通信链路；对传输中的分组进行路由选择和流量控制；传输数据的差错检测和纠正等。这些功能通常由数据链路层、网络层和传输层协议共同完成。

(2) 共享资源管理：采用有效的方法统一管理网络中的共享资源(硬件和软件)，协调各用户对共享资源的使用，使用户在访问远程共享资源时能像访问本地资源一样方便。系统为单机提供了进程管理、存储管理、文件系统管理和设备管理等共享功能。

(3) 网络管理：最基本的是安全管理，主要反映在用“存取控制”来确保数据的安全性，通过“容错技术”来保证系统故障时数据的安全性。此外，还包括对网络设备故障进行检测，对使用情况进行统计，以及为提高网络性能提供必要的信息。

(4) 网络服务：直接面向用户提供多种服务，如电子邮件服务、文件传输、存取和管理服务、共享硬件服务以及共享打印服务。

(5) 互操作：是指把若干相同或不同的设备和网络互连，用户可以透明地访问各服务点、主机，已实现更大范围的用户通信和资源共享。

(6) 网络接口：向用户提供一组方便有效的、统一的取得网络服务的接口，以改善用户界面，如命令接口、菜单、窗口等。

2. 网络操作系统的安全性

网络操作系统的安全性非常重要，主要表现在以下几个方面。

(1) 用户账号安全性：使用网络操作系统的每一个用户都有一个系统账号和有效的口令字。为了提高系统的安全性，必须在用户工作站发送口令字之前，对口令字加密。

(2) 时间限制：系统管理员对每个用户的注册时间进行限制，限制方式以一定的时间间隔为单位。时间限制功能主要应用在要求具有严格安全机制的网络环境中。

(3) 站点限制：系统管理员对每一用户注册的站点进行限制。站点限制了每个用户只能在指定物理地址的工作站上进行注册，以阻止企图从其他区域使用并进行注册的不良行为。

(4) 磁盘空间限制：系统服务员对每个用户允许使用的磁盘服务器磁盘空间加以限定，以防止可能出现的某些用户无限制侵占服务器磁盘的情况发生，确保其他用户磁盘空间的安全性。

(5) 传输介质的安全性：由于局域网的传输介质(如同轴电缆和双绞线)很容易被窃听，并将数据读走，因此网络传输介质的安全性也是十分重要的。从安全性考虑，网络传输介质应是光缆，因为对光缆的窃听非常困难。要在硬件上尽可能提供安全保障。

(6) 加密：对数据库和文件加密是保护文件服务器数据安全性的重要手段。一般在关闭文件时加密，在打开文件时解密。加密后具有超级用户特权的网络管理员才能读取服务器上的目录和文件。

(7) 审计：网络的审计功能可以帮助网络管理员对那些企图对网络操作系统实行窃听行为的用户进行鉴别。当对网络运行机理熟练的某用户通过多次重复敲入口令字来试探其他用户口令字时，很多网络就采取一定措施来制止这种非法行为。

11.1.2 Windows 操作系统

微软基于 NT 的操作系统主要有 Windows NT、Windows 2000、Windows XP、Windows Vista、Windows 2003 Server 和 Windows CE 等版本。

1. Windows NT 模型

微软推荐使用下列 4 种域模型中的任何一种：单个域、主域、多主域和完全信任域。

(1) 单个域模型：是最简单的 Windows NT 域模型。单域模型只使用一个域来为一个机构的用户和资源服务。它适用于几乎不关心安全问题的小型机构。

(2) 主域模型：用单域模型对用户账号信息实施控制，另外再用独立的资源域来管理。

(3) 多主域模型：它把两个或更多个主域以双向信任关系连接起来，管理许多资源域，就像在主域模型中一样，多主域模型中的资源域也是由许多独立的资源组成并且必须和主域有单向信任关系。

(4) 完全信任域模型：它与主域模型的不同之处在于前者的管理是完全分散的。该模型中的每一个域都管理它自己的用户、组、账号以及文件和打印共享信息。

2. Windows NT 主要支持的网络协议

(1) Net BEUI 协议：该协议是一种小型且快捷的协议，不太适合运用于较大型的网络。

(2) IPX/SPX 协议：它是 Novell Netware 的协议，在 Netware 的 LAN 上提供传输服务，支持中小型网络。IPX 对应于 OSI/RM 的网络层，SPX 对应于 OSI/RM 的传输层。

(3) TCP/IP 协议：是一个标准的、可路由选择的协议，是广域网和 Internet 访问的标准。

(4) DHCP 协议：DHCP 是 BOOTP 的扩展，它提供了一种动态指定 IP 地址和配置参数的机制，主要用于大型网络环境和配置比较困难的地方。

3. Windows NT 特点

(1) 具有兼容性和可靠性。Windows NT Server 的设计融入了对当今流行的应用环境，如 UNIX、OS/2 及 MS-DOS 等系统的支持。

(2) 内置强大的通信服务，如 TCP/IP、路由和远程访问等，都添加到嵌入式解决方案中。

(3) 完全支持 Win32 应用程序编程接口(API)，可以跨平台创建标准化应用程序。

(4) 具有高级编程性能，可以在一个面向对象的环境中快速地解决方案。

(5) 可以将 Microsoft 和第三方提供的管理特性集成到信息技术管理基础构架中。

(6) 是一个抢占式多任务、多线程操作系统，不同类型的应用程序可以同时运行。

4. 常用网络操作命令

在 PC“运行”中输入 cmd 后，进入命令行模式。常用网络操作命令如表 11.1 所示。

表 11.1 常用网络操作命令

命令	有关参数或举例说明	作　用
ping	-a(反向名字解析)、-t(显示统计信息)、-n(回声请求次数)、-l(回声请求报文)	检测网络连通性，用于预测故障和确定故障源
ipconfig	all、flushdns(刷新客户端 DNS 缓存的内容) displaydns(显示客户端 DNS 缓存的内容)	显示当前的所有网卡的 TCP/IP 配置参数
arp	-a (查询系统中缓存的 ARP 表) -a IP 地址如 arp-a 192.168.9.8	用于显示和修改地址解析协议(ARP)缓存表

续表

命令	有关参数或举例说明	作　用
tracert	tracert(检测故障的位置) tracert IP 地址(确定在哪个环节上出了问题) 如 tracert 192.168.0.9	跟踪数据包经过多少路由器可以到达目的地的路径
netstat	-s(显示每个协议的统计数据) -r(显示 IP 路由表的内容) -a(显示所有活动的 TCP 连接)	用于显示与 IP、TCP 等协议相关的统计数据,检验本机各端口的连接情况
route	-print(用于显示路由表中的当前项目) -add(将路由器项目添加给路由表) -add-p(用于初始化路由表)	用以显示、人工添加和修改路由表项目

11.1.3 其他操作系统

1. UNIX 网络操作系统

UNIX 操作系统是一个交互式的分时系统,提供了一个支持程序开发全过程的基础和环境。UNIX 操作系统通常被分为 3 个主要部分:内核(kernel)、shell 和文件管理。其中:

内核是 UNIX 操作系统的核心:直接控制着计算机的各种资源,能有效地管理硬件设备、内存空间和进程等,使用后程序不受错综复杂的硬件事件细节的影响。

shell 是 UNIX 内核与用户之间的接口,是 UNIX 的命令解释器。

文件管理是指对存储在存储设备(如硬盘)中的文件所进行的组织管理,通常是按照目录层次的方式进行组织。

UNIX 操作系统具有如下特征:

(1) 具有支持多个同时登录的用户的能力,是一个真正的多用户系统。

(2) 系统采用树形目录结构来组织各种文件及文件目录。

(3) 系统中文件是无结构的字节序列,外围设备如同磁盘上的普通文件一样被访问。

(4) 具有在后台开始进程的能力。

(5) 具有上百个子系统,其中包括几十种程序设计语言,200 多条命令对应着 200 个实用程序。如系统使用了灵活的命令程序设计语言 Shell,Shell 是一种程序设计语言。

(6) 良好的移植性,操作系统的可移植性带来了应用程序的可移植性,因而用户的应用程序既可用于小型机,又可用于其他的微型机或大型机。

(7) 用户定义的窗口系统中,最为流行的是 X-Windows 系统。

(8) 精巧的核心与丰富的实用层,UNIX 系统在结构上分成内核层和实用层。

2. Linux 网络操作系统

Linux 操作系统软件包不仅包括完整的操作系统,而且还包括了文本编辑器、高级语言编译器等应用软件。它还包括带有个窗口管理器的 X-Windows 图形用户界面,如同我们使用 Windows NT 一样,允许我们使用窗口、图标和菜单对系统进行操作。

Linux 与传统的网络操作系统相比,具有以下特点。

(1) 源代码公开:从诞生之日起,Linux 的源代码就是公开的,这是它与 UNIX、Windows NT 等传统网络操作系统最大的区别,使它一直能得到世界范围内的程序员来共享完善。

(2) 完全免费：Linux从内核到设备驱动程序、开放工具等，都遵从GPL协议，Internet上有大量关于Linux的网站和技术资料，可以免费下载，其中不包含任何有专利的代码。

(3) 适应多种硬件平台：Linux可运行的硬件平台较多，如IBM PC及其兼容机、Apple Macintosh计算机、Sun工作站等。

(4) 完全的多任务和多用户：Linux允许在同一时间内运行多个应用程序，允许多个用户同时使用主机。

(5) 稳定性好：运行Linux的服务器有公认的稳定性，很少出现在其他一些常用操作系统上常见的死机现象。

(6) 易于移植：Linux符合UNIX的标准，这使UNIX下的许多应用程序可以很容易地移植到Linux。

(7) 用户界面良好：Linux的X-Windows系统具有图形用户界面，它可以运行Windows 9X下的所有操作，甚至还可以在几种不同风格的窗口之间来回切换。

(8) 具有强大的网络功能：实际上，Linux就是依靠Internet才迅速发展起来的，它支持TCP/IP，支持网络文件系统、文件传送协议、超文本传送协议、点对点协议、电子邮件传送和接收协议和SMTP等，可以实现与其他网络操作系统互联。

11.2 网络操作系统与TCP/IP

11.2.1 应用程序接口软件API

NOS和OSI/RM中的分布大致如图11.1所示，从分层的角度讲，NOS主要包括3部分：网络驱动程序、网络协议软件和应用程序接口软件。

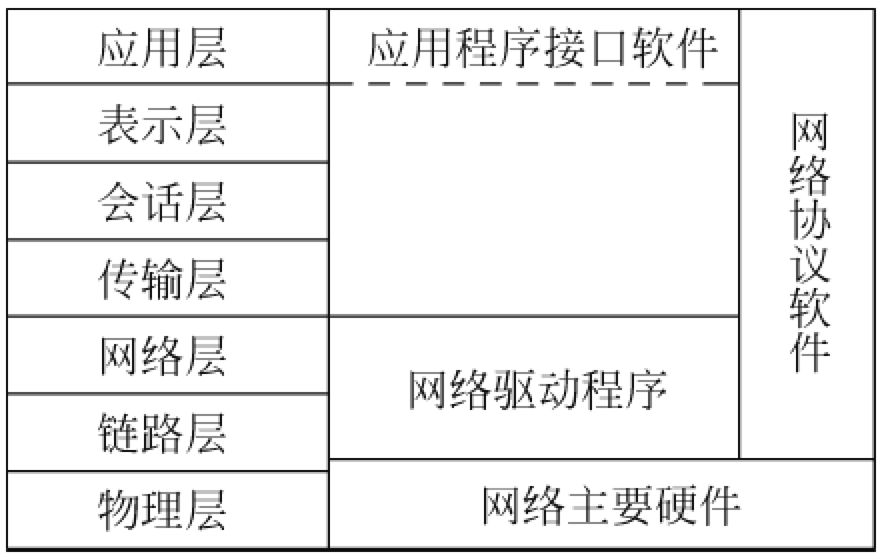


图11.1 NOS在OSI/RM中的分布

NOS几乎占据了OSI/RM对应的所有层，其中，网络驱动程序与网络主要硬件(分布于物理层和数据链路层)进行通信，驱动网络运行。例如，在局域网中网络驱动程序介于网络接口板(NIC)与网络协议软件之间，起中间联系作用。网络协议软件是在整个网络范围内传送数据单元所必需的通信协议软件。主要分布于OSI/RM的第2～7层。应用程序接口操作软件用于应用软件与网络协议软件的通信，支持NOS实现高层服务。下面介绍针对局域网的NOS分层结构。

1. 网络驱动程序

就局域网而言，网卡生产厂商必须提供每种网卡对应的驱动程序，以确保各种网卡都采用国际标准协议。通常，厂商随同网卡提供适用于不同操作系统的各种驱动程序。网络驱动程序屏蔽了网卡接收和发送数据单元的复杂处理过程，它直接对网卡的各种控制/状态寄存器、DMA(直接存储器存取)、I/O(输入/输出)端口进行硬件级操作。

2. 网络协议软件

由于网络协议软件几乎分布在网络的所有层，因此它直接关系到网络操作系统的性能。

例如，高速网络协议的软件会实现 NOS 的高速处理。

3. **应用程序接口软件(API)**

应用层提供多种应用协议和服务，其中应用服务与应用程序之间的接口软件完成本地系统与网络环境的联系。使用 TCP/IP 协议的应用程序通常会应用 API。

API 是应用程序和操作系统之间的接口软件。当某个应用进程启动系统调用时，控制权就从应用进程传递给了系统调用接口。此接口再将控制权传递给计算机的操作系统，操作系统将此调用给某个内部过程，并执行所请求的操作。内部过程一旦执行完毕，控制权就又通过系统调用接口返回给应用程序。只要应用进程需要从操作系统获得服务，就要将控制权传递给操作系统，操作系统在执行必要的操作后将控制权返回给应用进程，API 就是这种系统调用接口。

图 11.2 给出了系统调用、应用程序(包含系统命令以及其他应用程序)和 API 之间的关系。系统调用指操作系统提供给用户程序调用的一组“特殊”接口，用户程序可以通过这组“特殊”接口来获得操作系统内核提供的服务。系统调用并不直接与程序员进行交互，它仅仅是一个通过软件中断机制向内核提交请求以获取内核服务的接口。实际使用中，程序员通常只是用到用户编程接口(API)。系统为了更好地保护内核空间，将程序的运行空间分为用户空间和内核空间，也就是常说的的用户态和内核态，它们分别运行在不同的软件级别上，在逻辑上是相互分离的。

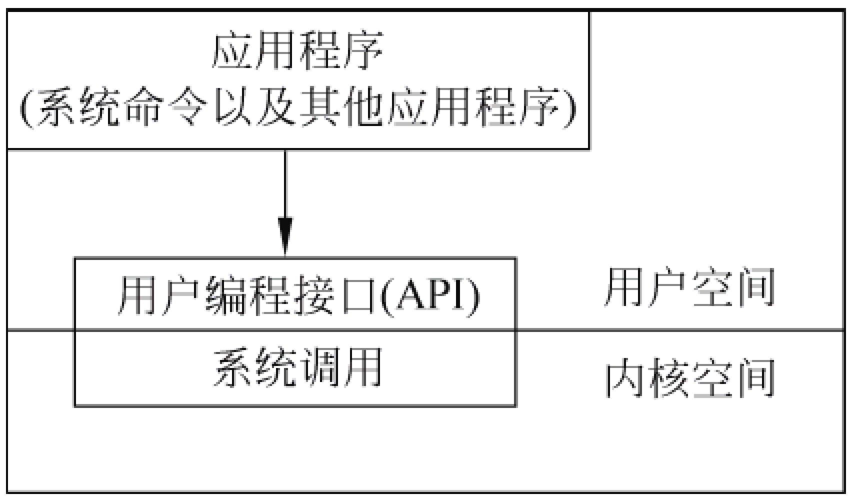


图 11.2 API 与其他系统之间的关系

11.2.2 Socket 基本函数

在 Windows 环境下运行的 Socket 基本函数，有 Bind(绑定套接字的本地端口号和本地 IP)、Listen(服务器随时能感知客户端的服务要求)、Accept(服务器接收客户端发来的连接请求)、Send(客户端和服务器之间发送数据)、recv(客户端和服务器之间接收数据)等函数。以下给出常用的几种系统 Socket 函数，函数括号中的参数项采用逗号隔离。

1. **WSAStartup 函数**

int WSAStartup(WORD wVersionRequested，LPWSADATA lpWSAData)

WSAStartup 为 Windows 异步套接字(Windows Sockets Asynchronous，WSA)的启动命令。

int 指套接字描述符是一个整数类型的值，每个进程空间中都有一个套接字描述符表。使用 Socket 的程序在使用 Socket 之前必须调用 WSAStartup 函数。

第 1 个参数：指明程序请求使用的 Socket 版本，高位字节为副版本、低位字节为主版本；

第 2 个参数：返回请求的 Socket 的版本信息。当一个应用程序调用 WSAStartup 函数时，操作系统根据请求的 Socket 版本来搜索相应的 Socket 库，绑定找到的 Socket 库到该应用程序中，以后应用程序就可以调用所请求的 Socket 库中的其他 Socket 函数了。

2. **WSACleanup 函数**

```
int WSACleanup (void)
```

应用程序在完成对请求的 Socket 库的使用后，要调用 WSACleanup 函数来解除与 Socket 库的绑定，并且释放 Socket 库所占用的系统资源。

3．socket 函数

```
SOCKET socket(int af, int type, int protocol)
```

系统应用程序调用 socket 函数来创建一个能够进行网络通信的套接字。socket 函数是用来创建指定套接字类型，并为套接字分配所需的系统资源。

第 1 个参数：用来指定使用的地址簇。对于 TCP/IP，该参数置为 AF_INET（或 PF_INET）；

第 2 个参数：用来指定创建的套接字类型，流式套接字为 SOCK_STREAM，数据报套接字为 SOCK_DGRAM，原始套接字为 SOCK_RAW；

第 3 个参数：依赖第 2 个参数，指定应用程序所使用的通信协议。如果指定为 0，那么系统就会根据地址格式和套接类别，自动选择一个合适的协议。IP 协议为 IPPROTO_IP，TCP 协议为 IPPROTO_TCP，UDP 协议为 IPPROTO_UDP，通常设为 IPPROTO_IP。

如果函数执行成功返回一个新的套接字描述符，失败返回 INVALID_SOCKET 错误。

每个进程的进程空间里都有一个套接字描述符表，该表中存放着套接字描述符和套接字数据结构的对应关系。该表中有一个字段存放新创建的套接字的描述符，另一个字段存放套接字数据结构的地址，因此根据套接字描述符就可以找到其对应的套接字数据结构，数据结构存放在操作系统的内核缓冲区中。

4．accept 函数

```
SOCKET accept(SOCKET s, struct sockaddr FAR * addr, int FAR * addrlen)
```

accept 函数用来接收客户端的连接建立请求。

第 1 个参数：用来指定服务器接收请求的流套接字，该套接字已通过 listen 函数设置为倾听状态；

第 2 个参数：用来指定接收请求的套接字地址（保存客户端 IP 地址和端口）；

第 3 个参数：用来指定套接字长度。

服务程序调用 accept 函数中处于监听状态的流套接字 s 的队列里排在最前的一个客户连接请求后，创建一个新的套接字来与客户端套接字构成连接通道，如果连接成功，就返回新创建的套接字的描述符，以后与客户套接字交换数据的是新创建的套接字。

5．bind 函数

```
int bind(SOCKET s, const struct sockaddr FAR * name, int namelen)
```

bind 函数用来将套接字与本地地址相互绑定。由于当创建了一个 Socket 以后，套接字数据结构中有一个默认的 IP 地址和默认的端口号，一个服务程序必须调用 bind 函数来给其绑定一个 IP 地址和一个特定的端口号。

第 1 个参数：指定待绑定的 Socket 描述符；

第 2 个参数：指定一个 sockaddr 结构长度。

6．connect 函数

```
int connect(SOCKET s, const struct sockaddr FAR * name, int namelen)
```

这个函数专为流式套接字设计，为请求与服务器建立连接，用于面向连接的 TCP 类型的服务。如此理解 connect 三个参数就容易了，首先必须指定数据发送的地址，同时也必需指定数据从哪里发送，这正好是 connect 的前两个参数，而第三个参数是为第二个参数服务的。

第 1 个参数：指定数据发送的套接字，解决从哪里发送的问题；

第 2 个参数：指定数据发送的目的地，也就是服务器端的地址。服务器是被动连接的一方需要调用 listen 以接收 connect 的连接请求；

第 3 个参数：指定 sockaddr 结构体的长度。

客户程序调用 connect 函数来使客户 Socket s 与监听于 name 所指定的计算机的特定端口上的服务 Socket 进行连接。

7. closesocket 函数

```
int closesocket(SOCKET s)
```

执行此函数，操作系统会关闭 s 套接字，释放相应资源。

8. listen 函数

```
int listen(SOCKET s, int backlog)
```

listen 函数用来监听端口上的连接建立请求，使流式套接字 s 处于监听状态。这个函数专为流式套接字设计，用于面向连接的 TCP 类型的服务。s 用来指定服务器端要监听的套接字；backlog 用来指定流式套接字要维护的客户连接请求队列的最大长度，如果 backlog 设置为 SOMAXCONN，下层的服务提供者会将客户连接请求队列设置为最大的合理值。

9. send 函数

不论是客户还是服务器应用程序都用 send 函数来向 TCP 连接的另一端发送数据。客户程序一般用 send 函数向服务器发送请求，而服务器则通常用 send 函数来向客户程序发送应答。

```
int send(SOCKET s, const char FAR * buf, int len, int flags)
int sendto(SOCKET s, const char FAR * buf, int len, int flags, const char FAR * to, int tolen)
```

send 函数是专门为流式套接字设计，用于面向连接的 TCP 类型的服务；

sendto 函数专门为数据报套接字设计，用于无连接的 UDP 类型服务(共 6 个参数)。

第 1 个参数：指定发送端套接字描述符；

第 2 个参数：指明一个存放应用程序要发送数据的缓冲区；

第 3 个参数：指明实际要发送的数据的字节数；

第 4 个参数：一般置 0。这里只描述同步 Socket 的 send 函数的执行流程；

第 5 个参数：用来指定接收端保存等待接收数据的缓冲区；

第 6 个参数：用来指定接收数据的字节数。

注意：并不是 send 把 SOCKET s 的发送缓冲区的数据传到连接的另一端的，而是由相关协议传的，send 仅仅是把 const char FAR * buf 中的数据复制到 SOCKET s 的发送缓冲区的剩余空间里。

10. recv 函数

不论是客户还是服务器应用程序都用 recv 函数从 TCP 连接的另一端接收数据。

```
int recv(SOCKET s,char FAR * buf,int len,int flags)
recvfrom(SOCKET s,char FAR * buf,int len,int flags,const char FAR * from,int * fromlen)
```

recv 函数专门为流式套接字设计,用于面向连接的 TCP 类型服务;
recvfrom 函数为数据报套接字设计,用于无连接的 UDP 类型服务(共 6 个参数)。
第 1 个参数:指定接收端套接字描述符;
第 2 个参数:指明一个缓冲区,该缓冲区用来存放 recv 函数接收到的数据;
第 3 个参数:用来指定 buf 接收数据的字节数;
第 4 个参数:flags 是附加标志,一般置 0;
第 5 个参数:用来指定发送端保存等待发送数据的缓冲区;
第 6 个参数:用来指定发送数据的字节数。
注意:recv 函数仅仅是 SOCKET s 中收到的数据 copy char FAR * buf 接收缓冲区。

11.2.3 C/S 构架下的 Socket 通信

1. C/S 模式构架

网络操作系统通过网络实现互相传递数据与各种消息的功能,可分为服务器(Server)及客户端(Client)。服务器的主要功能是管理服务器和网络上的各种资源和网络设备的共用,加以统合并控管流量,避免网络瘫痪的可能性,而客户端就有着能接收服务器所传递的数据来运用的功能,以便客户端可以清楚地搜索所需的资源。

为降低开销,按计划将任务分配到服务器和客户端,通过安装客户端来进行相应的管理操作。服务器端和客户端的程序不同,用户所需要的各种功能的程序主要放在客户端,以完成用户的具体的业务,服务器端主要提供数据管理、传输管理和数据分享等业务。

C/S 模式结构如图 11.3 所示。客户端的应用程序、中间件和服务器端管理程序 3 个部分是 C/S 模式的基本结构。服务器端一般分为应用和数据库两种服务器。C/S 结构的实现原理是:Client 端首先通过用户的操作形成相应的语句,然后通过 TCP/IP 的网络协议向服务器发送相应的命令语句;服务器端通过对应监听端口对服务请求进行实时检测。服务器端一旦发现请求,先验证客户的身份,在验证通过之后才开始执行请求客户的相应命令。

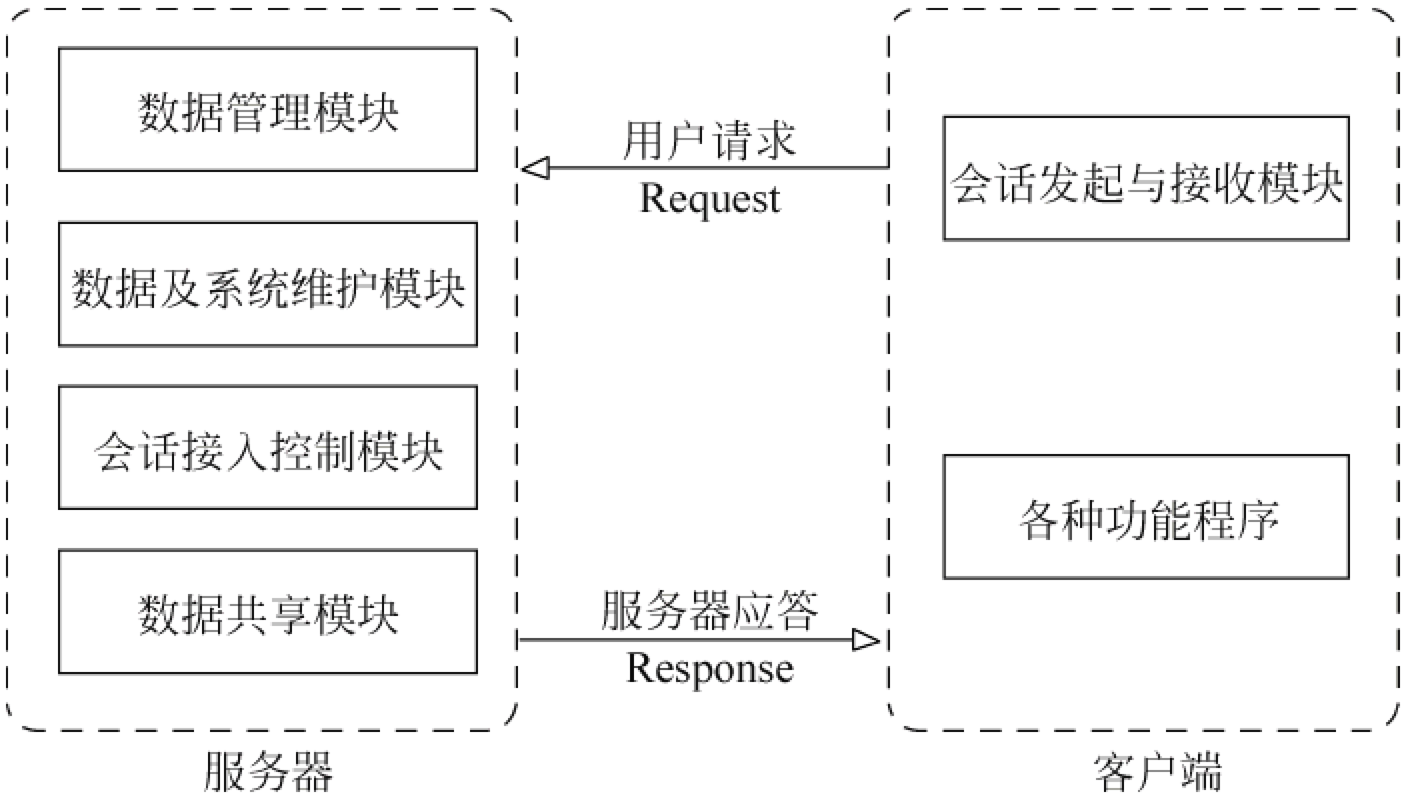


图 11.3 C/S 模式构架图

2. Socket 通信模型

在 C/S 模式中基于 TCP 协议使用套接字(Socket)的通信模型如图 11.4 所示。Socket 完成通信的步骤是：在服务器端先创建 Server Socket 对象并绑定监听端口，然后调用 accept 函数来监听客户端的请求，此时若接到客户端实例化的 Socket 类发来的请求后便打开连接通道，通过输入流读取客户端发来的请求信息，通过输出流向客户端发送（或转发）信息。此时客户端的信息交互方式与服务器类似，通过输入流接收服务器输出流发来的消息，通过输出流向服务器的输入流发送消息，在图中表示为调用 InputStream 和 OutputStream 中的类进行通信传输；在流传输结束后双方都会关闭输入、输出流并关闭 Socket，此时服务器的 ServerSocket 类依旧运行 accept 函数，等待客户端传来新的连接请求。

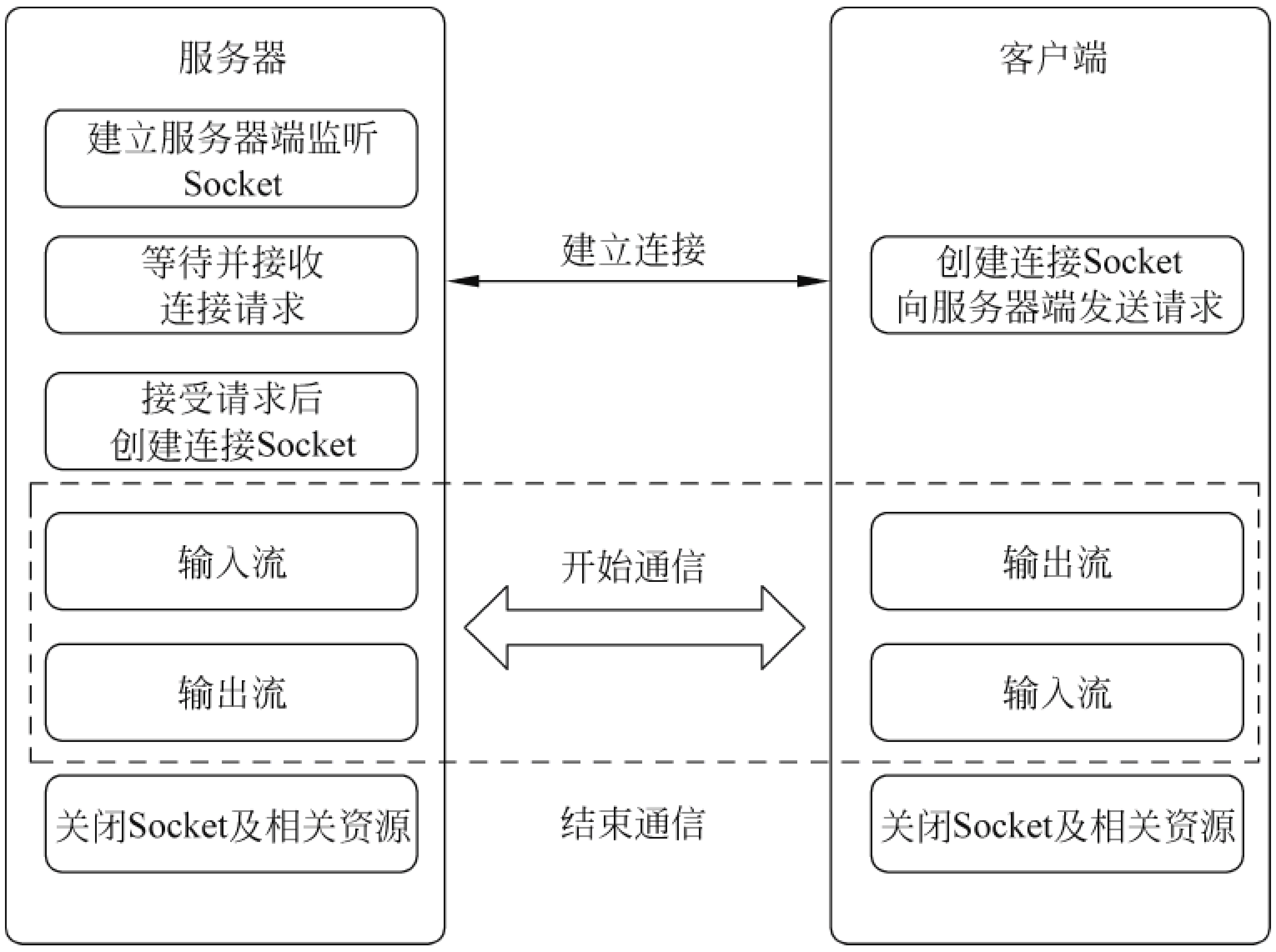


图 11.4 Socket 通信模型

使用 TCP/IP 协议的应用程序通常会采用 API，在网络上的两个程序通过一个双向的通信连接来实现对数据的交换，这个双向链路的两端就是 Socket，多用来实现客户端和服务器之间的数据连接，是 TCP/IP 协议中的一个很常用的编程界面，而对其进行具体化则需要采用 Socket 接口来进行实现，实际上 Socket 是对 TCP/IP 协议的封装，Socket 本身并不是协议，而是一个调用接口，也就是一个 API，通过 Socket 才能使用 TCP/IP 协议。

TCP/IP 只是一个协议栈，就像操作系统的运行机制一样，必须要具体实现，同时还要提供对外的操作接口。这个就像操作系统会给程序员提供标准的编程接口，如 Win32 编程接口。TCP/IP 提供给程序员的编程接口就是 Socket。一对 IP 地址和端口号就可以唯一确定一个 Socket 连接，这个连接就是服务器和客户端的通信。IP 地址相当于发端知道了通信对方办公的地址，但一个办公室有好多工位，而具体工位就相当于端口号。

3. Socket 连接过程

Socket 之间的连接过程分为三个步骤：服务器监听，客户端请求，连接确认。

(1) 服务器监听：服务器端套接字并不定位具体的客户端套接字，而是处于等待连接的状态，实时监控网络状态，等待客户端的连接请求。

(2) 客户端请求：指客户端的套接字提出连接请求，要连接的目标是服务器端的套接

字。为此,客户端的套接字必须首先描述它要连接的服务器的套接字,指出服务器端套接字的地址和端口号,然后就向服务器端套接字提出连接请求。

(3) 连接确认：当服务器端套接字监听到或者说接收到客户端套接字的连接请求时,就响应客户端套接字的请求,建立一个新的线程,把服务器端套接字的描述发给客户端,一旦客户端确认了此描述,双方就正式建立连接。而服务器端套接字继续处于监听状态,继续接收其他客户端套接字的连接请求。

4. Socket 的应用

实例：在 C/S 模式下实现基于 TCP 协议的 Socket 通信联系,要求通过服务器中间转发,完成不同客户端之间的信息交流。

首先服务器和客户端都需要调用 Socket 接口,其中服务器是被动连接的一方需要调用 listen 函数以接收 connect 的连接请求,对客户端的接入进行监听;而客户端则是通过建立包含服务器的 IP 地址以及端口信息在内的 Socket 接口,通过调用 WSAStartup 函数来完成接入信号的发出。

处于监听状态的服务器一旦接收到客户端发来的接入请求后,就会建立起应用层的 Socket 虚连接(传输层采用 TCP 协议)。然后客户端以字符输入为流式套接字的形式向传输信道中写入自己的登录目的客户端的信息,信息中包含登录名、IP 地址、接入端口号等,以供服务器进行检测判断和转发。

此时的服务器从传输信道中读取到客户端发来的登录信息并进行检查。若满足接入条件要求(例如在线人数符合上限人数、未出现用户信息重复等情况),就会在服务器为它建立 Socket 连接表,连接表如表 11.2 所示。该表包含登录名、密码、IP 地址、端口号以及客户端用户的基本信息。否则将会断开之前建立的 Socket 连接。

表 11.2 Socket 连接表

序号	登录名	密码	IP 地址	端口号	其他信息
1	ABC	123456	22.33.44.55	8088	abc
…					

服务器和客户端建立的连接成功后,便向该客户端回传成功登录信息,并向其传输当前的在线用户列表,以方便该用户与其他用户进行通信。

服务器再向其他已经与其建立连接的客户端发送当前用户登录信息,以使其他客户端更新在线用户列表,在此之后可以进行后续的数据传输。由于传输层采用的是 TCP 面向连接的协议,所以连接状态是一直都在维持的,即使此时并无数据传输。

基于 TCP 传输协议的 Socket 通信中,客户端之间在应用层建立起了一条数据传输的逻辑通道。客户端要传输的信息应该具有以下三要素：

(1) 目的客户端的信息(如登录名等)和源客户端自己的信息。

(2) 信息类型的关键字,如文字信息、图片信息和语音信息等,然后设定关键字来分别表示这几种类型的信息。

(3) 信息具体内容,如发送某一张图片或聊天文字等。

服务器从收到的数据报中提取这三个要素后,首先要判断信息的去向,将目的客户端信息与 Socket 连接表里的信息进行比较,以确定目的客户端的 IP 地址和端口号;接着提取关

键字，对信息类型进行判断；然后按类型通读取信息内容，如对文字类型信息以字符流方式传输；对图片、语音类型以字节流方式传输。

下一步的工作就是服务器将读取的信息转发给目的客户端，当然转发的数据报也要具备三个要素：信息来源客户端的信息（登录名）、要传达的信息种类的关键字和信息内容。

目的客户端在收到报文后，首先将源客户端信息进行输出显示；随后进行信息类型关键字的判断，然后通过合适的方式读取信息内容并对内容进行显示，让目的客户端的用户可以获取相应的信息。

至此客户端之间通过服务器转发而进行的信息交流的流程就已经完成，在经过一段时间的交流以后客户端要结束通信，即断开连接，此时需要客户端通过已经建立的 Socket 连接向服务器发送退出登出信息，其中包含源用户信息以及登出信号关键字，服务器在接收到该信息确认无误后，向客户端传回确认断开信息，并关闭 Socket 连接。该客户端接收到确认信息后也断开 Socket 连接，使得这次通信告终。

11.3 Windows Server 2016 安装

Windows Server 是微软最早在 2003 年推出的 Windows 的服务器操作系统，现在用得最多的是 Windows Server 2012，下面要介绍的是最新版本 Windows Server 2016 在 Oracle VM VirtualBox 虚拟机中的安装过程。

VirtualBox 和 Windows Server 2016 可以在下载后再安装。

11.3.1 在 VirtualBox 中配置虚拟机

VirtualBox 管理器界面如图 11.5 所示，以下是在 VirtualBox 中配置一台虚拟机的过程。

图 11.5　VirtualBox 中配置虚拟机

(1) 打开 Virtual 虚拟机软件,创建新的虚拟机。

(2) 在名称中输入“Windows Server 2016”,类型选择“Microsoft Windows”,版本选择“Windows 2016(64-bit)”。

(3) 选择分配给虚拟计算机的内存大小,一般为计算机物理内存的一半。

(4) 创建虚拟硬盘,并设置虚拟硬盘类型为 VHD(虚拟硬盘),文件分配方式为动态分配,文件大小设置为 32GB 以上。

(5) 启动虚拟机。选择 Windows Server 2016 虚拟光盘文件或已放入 Windows Server 2016 光盘的光驱为启动盘。

11.3.2 Windows Server 2016 的安装

开始安装后需要首先选择语言、时间和货币格式以及键盘和输入方法,如图 11.6 所示。

图 11.6 选择语言、时间和货币格式以及键盘和输入方法

接下来选择安装的版本,选择安装版本如图 11.7 所示。共有 4 个版本可供选择:标准评估版、标准评估版(桌面体验)、数据中心评估版和数据中心评估版(桌面体验)。如果习惯了 Windows 桌面,那么一定要选择带桌面体验的版本。

接下来需要选择安装的类型,如图 11.8 所示。如果机器上原本就有 Windows Server 的旧版本,可以选择“升级:安装 Windows 并保留文件、设置和应用程序”。如果是从头安装 Windows Server,可以选择“自定义:仅安装 Windows(高级)(C)”。在这里选择了后一种类型。

下面需要选择安装的位置,如图 11.9 所示。在磁盘分区列表中选择要安装的位置。如果磁盘未分区,也可以直接选择“下一步”。

在安装的最后一步,需要设置内置管理员账户“Administrator”的密码,如图 11.10 所示。设置密码时需用大小写字母、标点符号或数字这 4 类字符中的 3 类的组合。

单击“完成”按钮。安装完成后重启进入管理员登录界面。

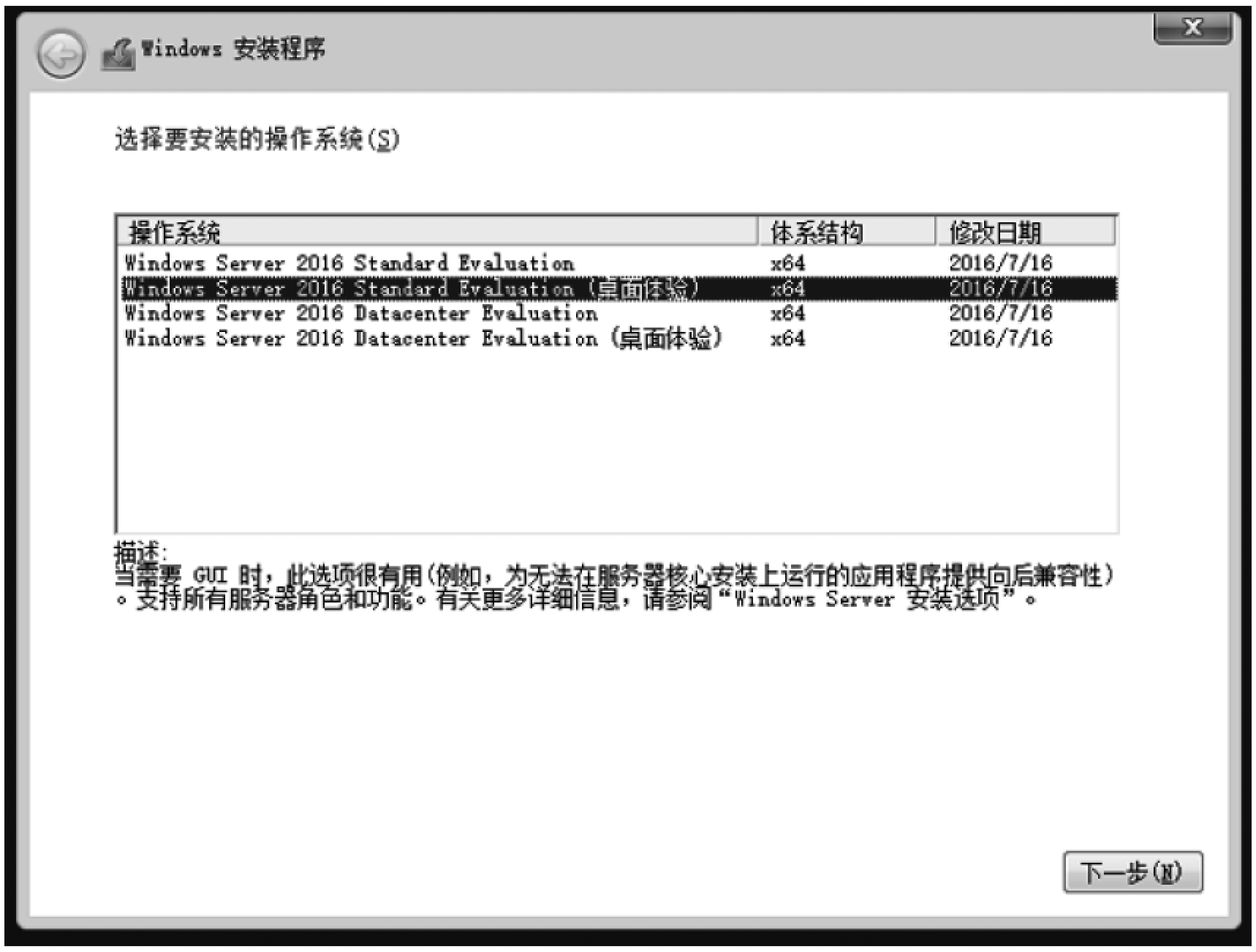


图 11.7　选择安装版本

图 11.8　选择安装类型

图 11.9　选择安装位置

图 11.10　设置内置管理员密码

登录后单击“开始”按钮，弹出“开始”菜单。选择“设置”菜单项，打开 Windows 设置窗口，菜单项如图 11.11 所示。

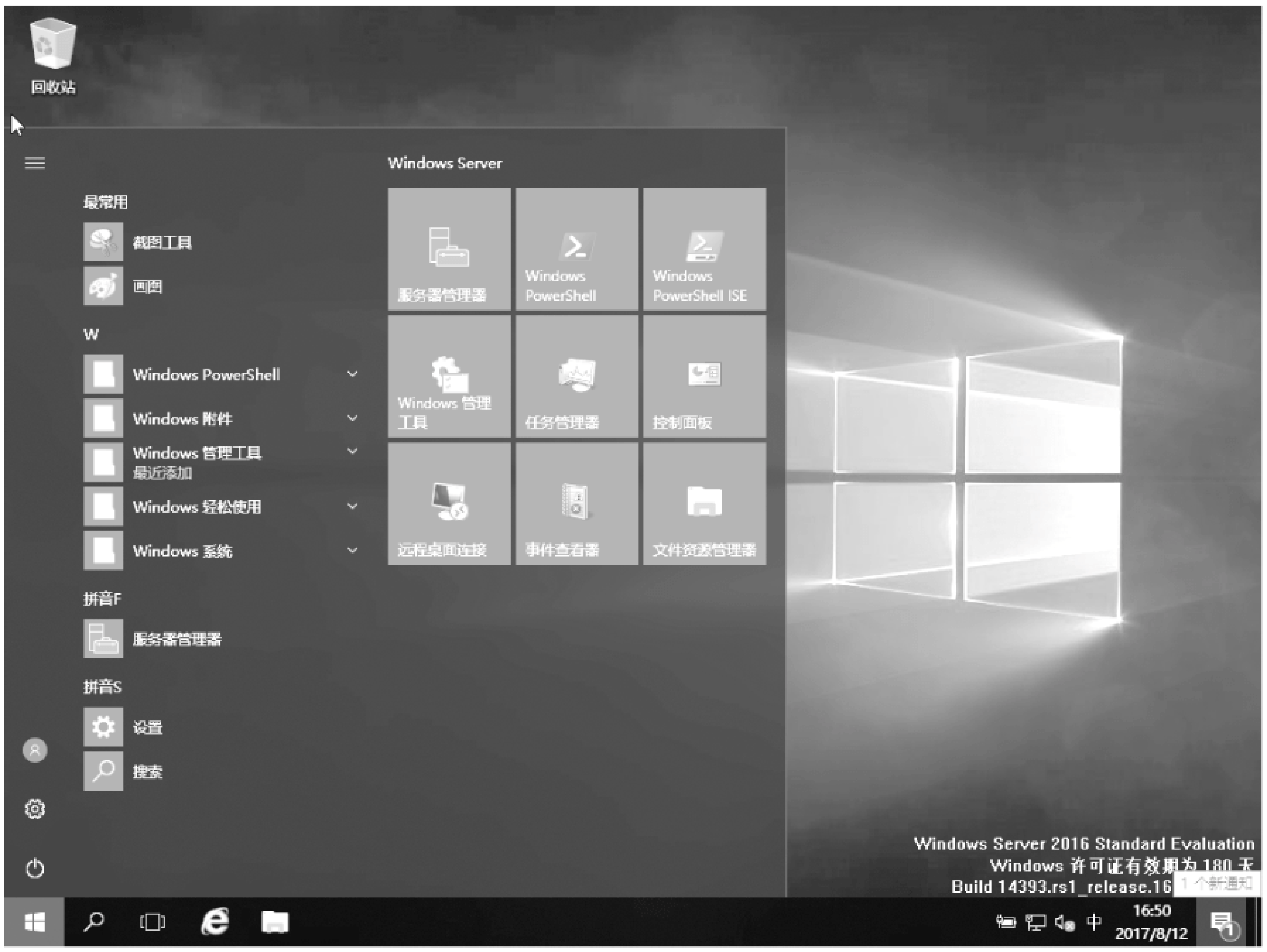


图 11.11　开始菜单

进入图 11.12 后，单击“网络和 Internet”，打开网络和 Internet 设置窗口，如图 11.13 所示。

图 11.12　Windows 设置

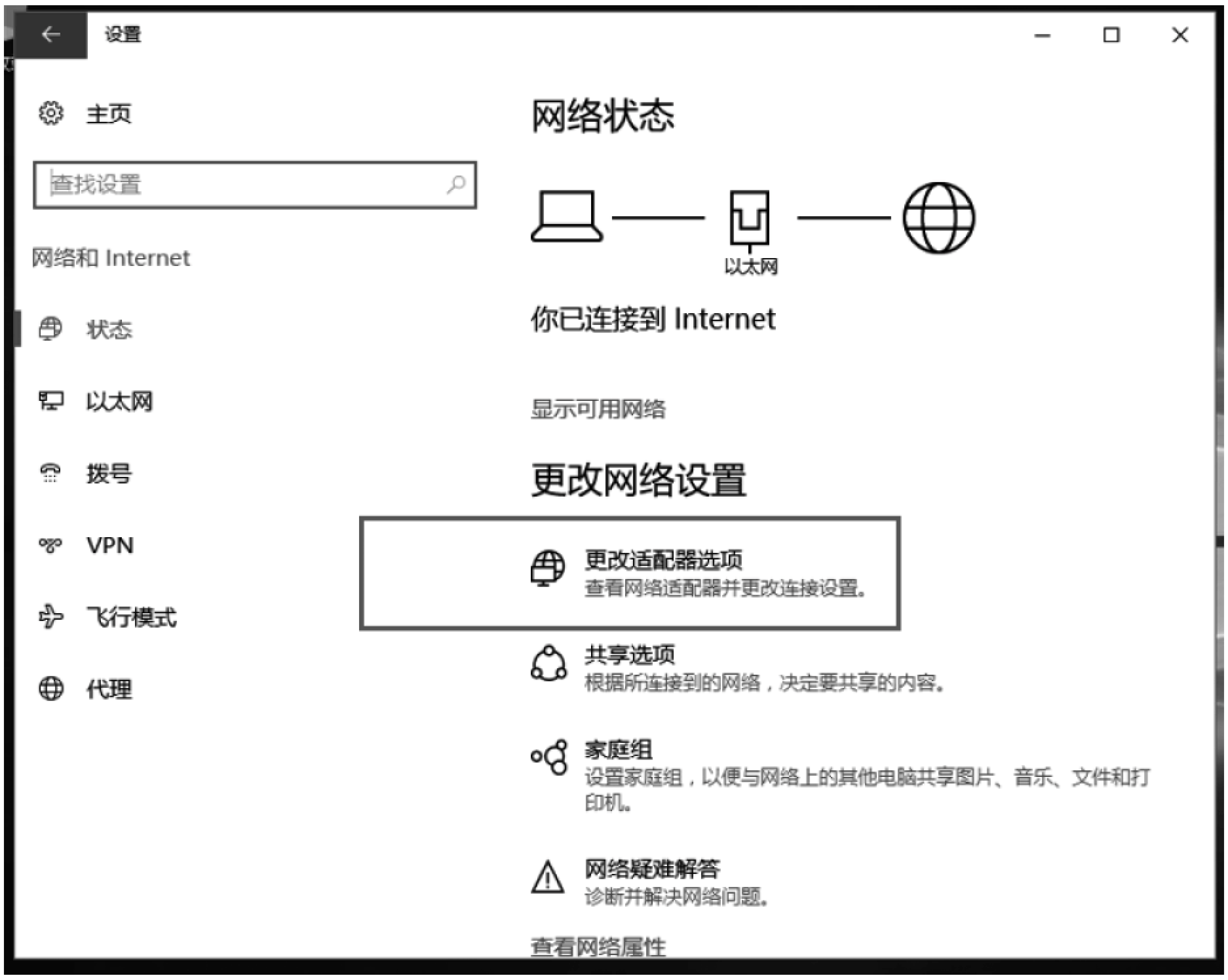


图 11.13　网络和 Internet 设置

在图 11.13 中，单击“更改适配器选项”，打开网络连接窗口，如图 11.14 所示。

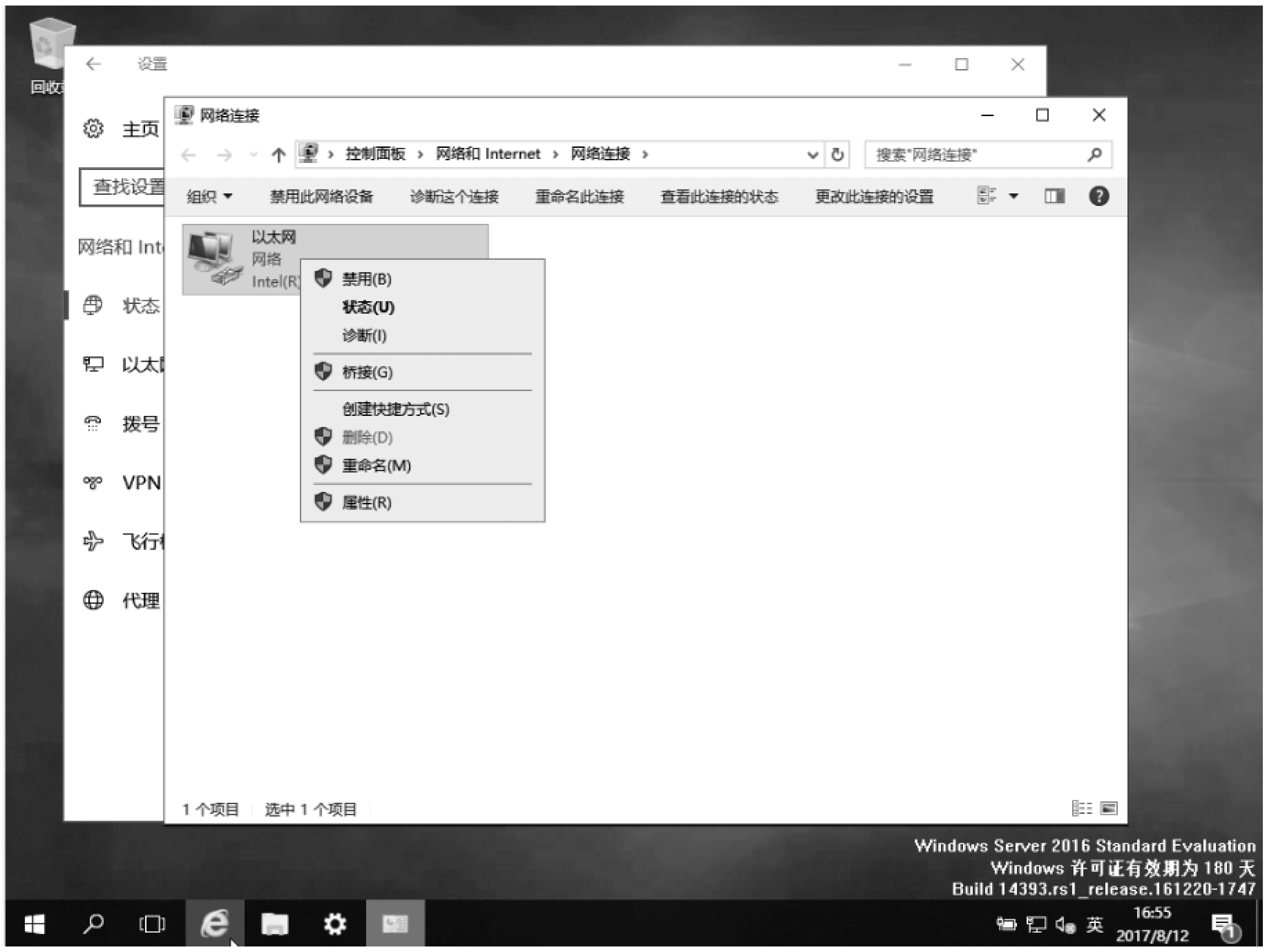


图 11.14 网络连接设置

选择所用的网络连接，单击右键，在快捷菜单中选择“属性”，弹出网络连接属性窗口，如图 11.15 所示。

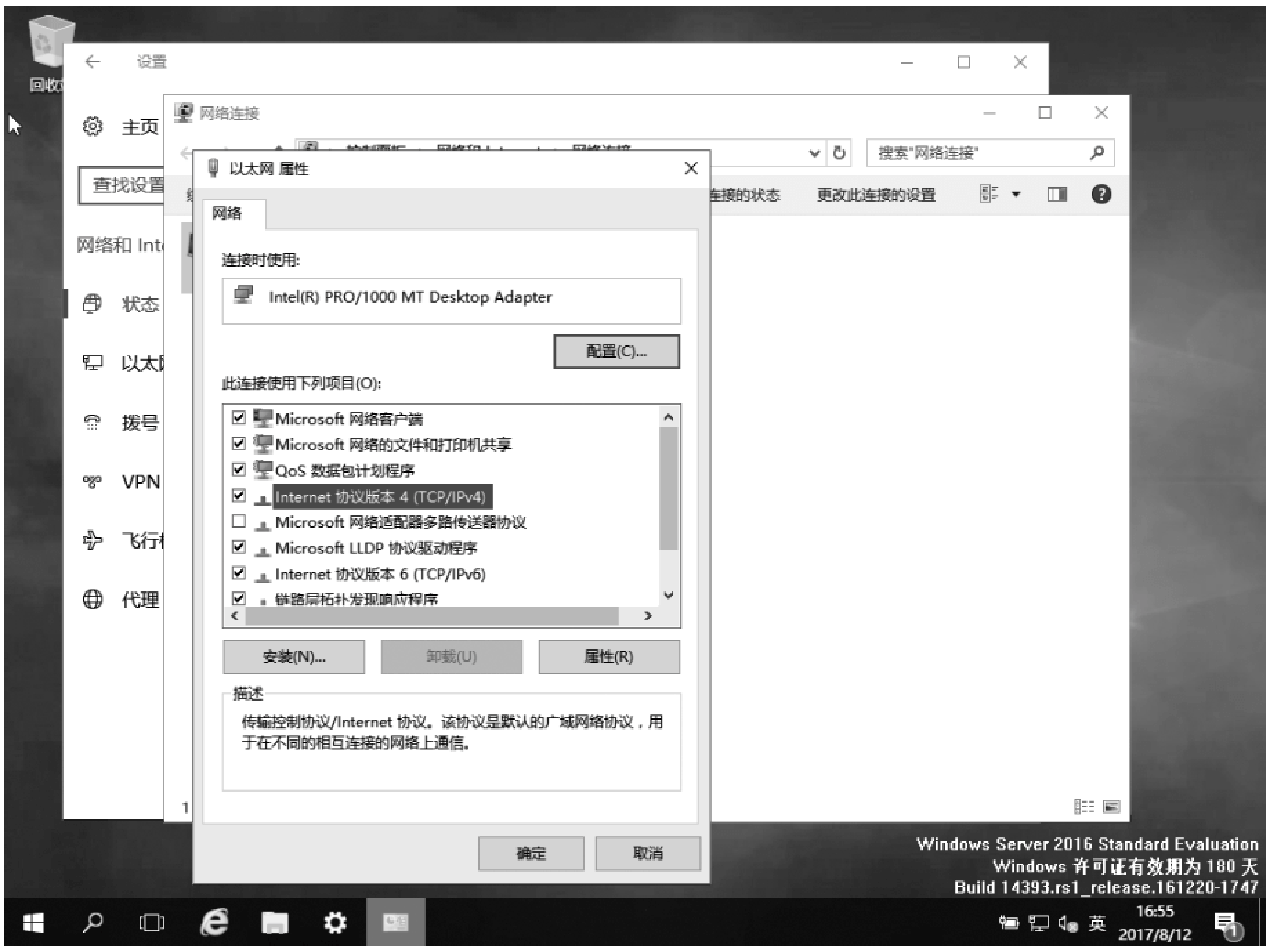


图 11.15 网络连接属性

在图 11.15 所示的网络连接属性窗口中，选择“Internet 协议版本 4（TCP/IPv4）”，打开 Internet 协议版本 4（TCP/IPv4）属性窗口，如图 11.16 所示。当然本机局域网采用的是 IPv6 地址，则选择“Internet 协议版本 6（TCP/IPv6）”，可打开 Internet 协议版本 6（TCP/IPv6）属性窗口。

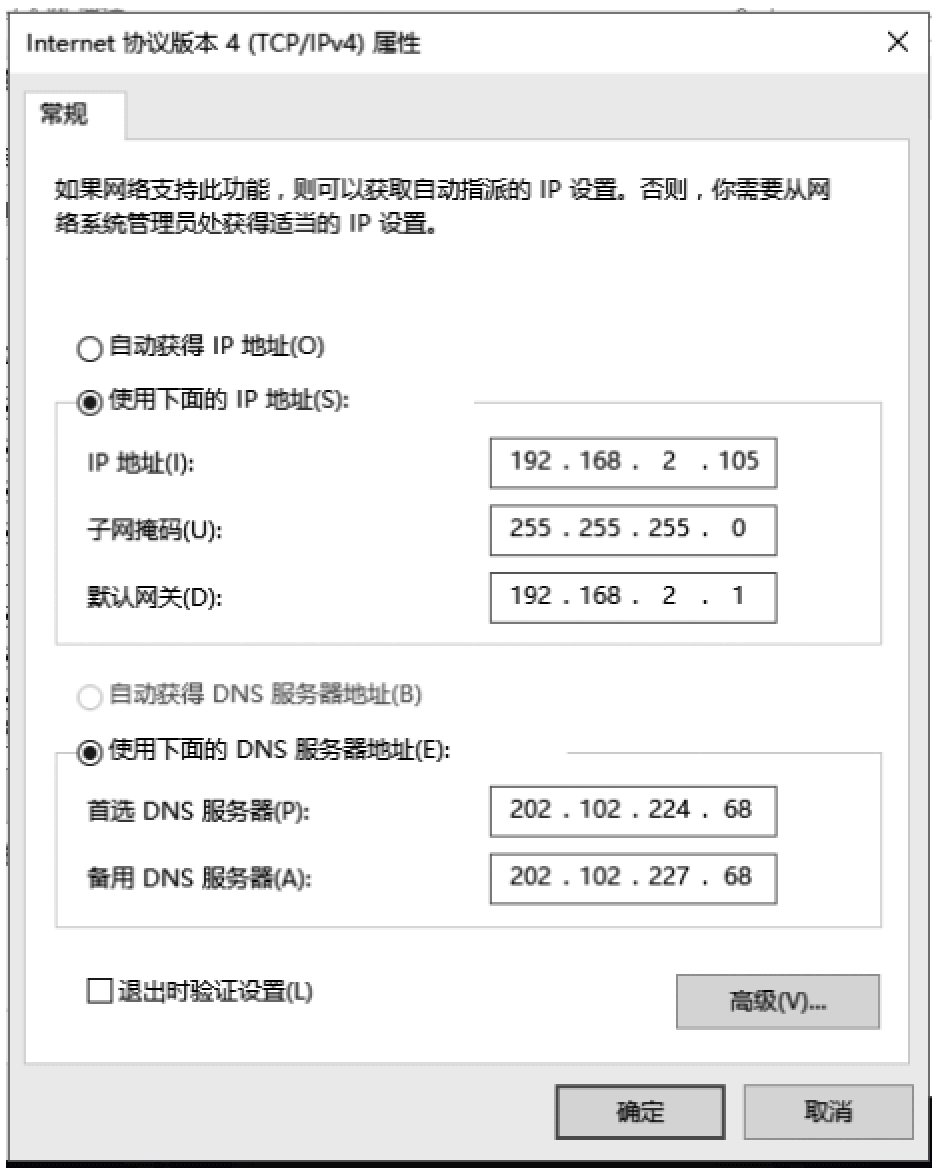


图 11.16　TCP/IPv4 属性

最后设置虚拟机的网卡连接方式。VirtualBox 提供了 4 种网卡连接方式：

（1）NAT 网络地址转换模式（Network Address Translation，NAT）。

（2）Bridged Adapter 桥接网卡模式。

（3）Internal 内部网络模式。

（4）Host-only Adapter 主机模式。

NAT 是默认方式。桥接网卡方式基本上和主机一样，也是一种常用的方式。它通过主机网卡架设了一座桥，直接连入到网络中。因此，它使得虚拟机能被分配到一个网络中独立的 IP，所有网络功能完全和网络中的真实机器一样。在此将网卡连接方式设置为桥接网卡方式，如图 11.17 所示。具体步骤：单击虚拟机软件的“控制”→“设置”菜单，打开设置窗口。选择“网络”，将网卡连接方式选择为“桥接网卡”。

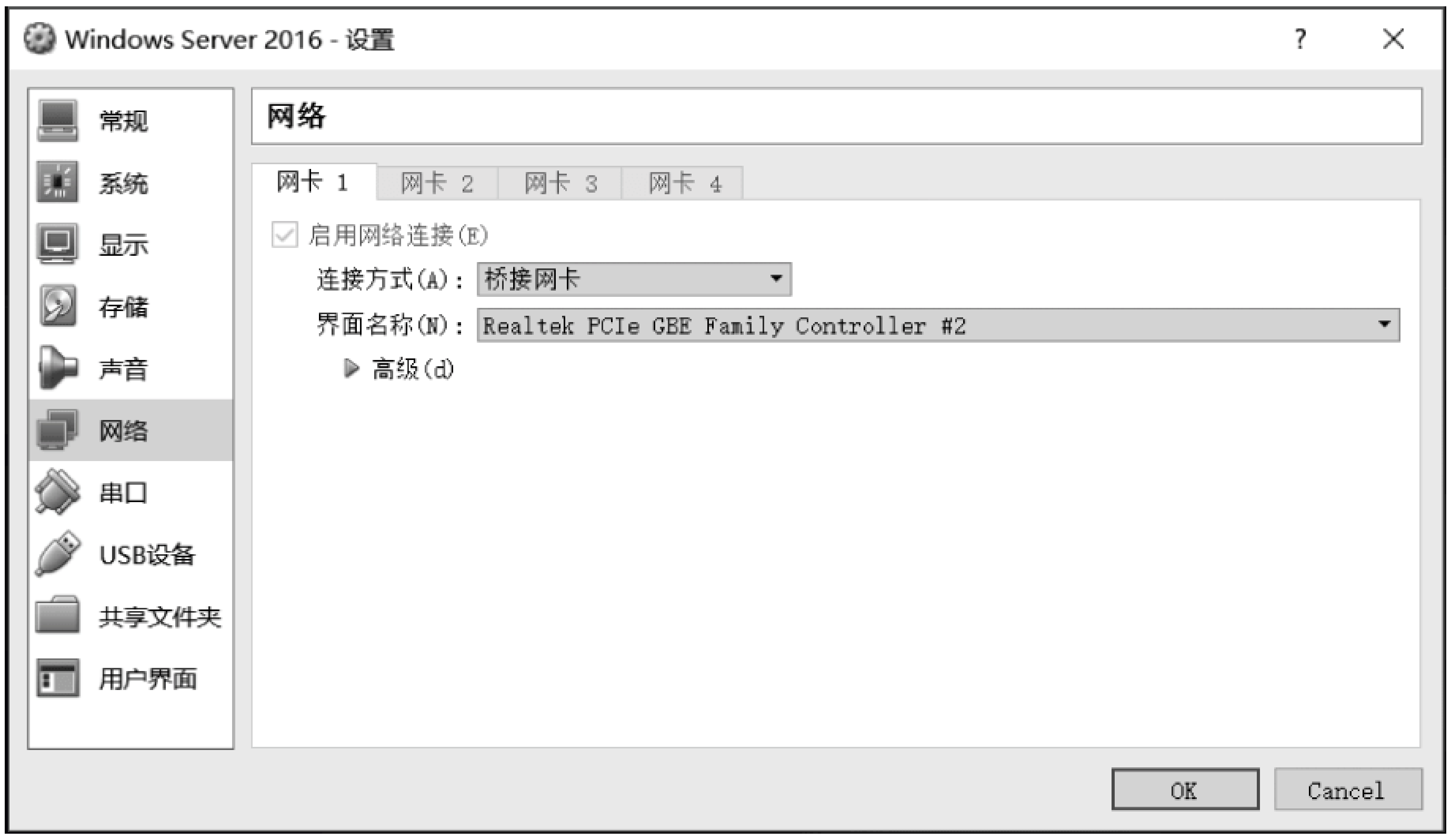


图 11.17　设置虚拟机网卡连接方式

11.4　Windows 服务器配置

Windows Server 配置有数十项之多，本小节主要讲述 DNS、DHCP、Web 服务、FTP 等的配置。

11.4.1　服务器角色选择

在该机器上配置服务器，可在“开始”菜单中选择“服务器管理器”菜单项，如图 11.18 所示，打开“服务器管理器”窗口，在左侧树形导航控件选择仪表板（默认选项）。

图 11.18　服务器管理器

单击图 11.19 中“添加角色和功能”，将启动向导，为服务器添加角色和功能。

图 11.19　添加角色和功能向导

接下来，需要选择安装类型。选择“基于角色或基于功能的安装”，选择安装类型如图 11.20 所示，单击“下一步”按钮。

图 11.20　选择安装类型

接着选择要安装角色和功能的目标服务器或虚拟硬盘。服务器选择如图 11.21 所示。

图 11.21　服务器选择

单击“下一步”按钮，选择要安装在服务器上的一个或多个角色。服务器角色选择如图 11.22 所示。再根据向导提示，添加所选的角色，进行安装。安装成功后角色或功能会出现在服务器管理器的仪表板中。

图 11.22　服务器角色选择

11.4.2　Web 服务器配置

Microsoft 的 IIS(Internet Information Services)是允许在公共 Intranet 或 Internet 上发布信息的 Web 服务器。IIS 是目前最流行的 Web 服务器产品之一，很多著名的网站都是建立在 IIS 的平台上。IIS 提供了一个图形界面的管理工具，称为 Internet 服务管理器，可用于监视配置和控制 Internet 服务。

IIS 是一种 Web 服务组件，其中包括 Web 服务器、FTP 服务器、NNTP 服务器和 SMTP 服务器，分别用于网页浏览、文件传输、新闻服务和邮件发送等方面，它使得在网络(包括 Internet 和局域网)上发布信息成了一件很容易的事。它提供 ISAPI(Intranet Server API)作为扩展 Web 服务器功能的编程接口；同时，它还提供一个 Internet 数据库连接器，可以实现对数据库的查询和更新。

一般在安装操作系统时不默认安装 IIS，所以在第一次配置 Web 服务器时需要安装 IIS。IIS 的安装方法为：在服务器管理器中启动添加服务器角色向导。在图 11.22 所示的角色选择中，选择“Web 服务器(IIS)”，并单击“添加功能”，选择服务器所需功能，执行安装。安装完成后，服务器管理器仪表板中将出现 IIS 模块，如图 11.23 所示。

图 11.23　Web 服务器配置完成

11.4.3　DNS 服务器配置

DNS(域名系统)是由解析器和域名服务器组成的。域名服务器是指保存有该网络中所有主机的域名和对应 IP 地址,并具有将域名转换为 IP 地址功能的服务器。其中域名必须对应一个 IP 地址,而 IP 地址不一定有域名。域名服务器为 C/S 模式中的服务器方,它主要有两种形式:主服务器和转发服务器。将域名映射为 IP 地址的过程就称为“域名解析”。在 Internet 上域名与 IP 地址之间是一对一(或者多对一)的,也可采用 DNS 轮循实现一对多,域名虽然便于人们记忆,但机器之间只认 IP 地址,它们之间的转换工作称为域名解析,域名解析需要由专门的域名解析服务器来完成,DNS 就是进行域名解析的服务器。DNS 命名用于 Internet 等 TCP/IP 网络中,通过用户友好的名称查找计算机和服务。

DNS 是一种组织成层次结构的分布式数据库,里面包含有从 DNS 域名到各种数据类型(如 IP 地址)的映射。这通常需要建立一种 A(Address)记录,意为“主机记录”或“主机地址记录”,是所有 DNS 记录中最常见的一种。通过 DNS,用户可以使用友好的名称查找计算机和服务在网络上的位置。DNS 名称分为多个部分,各部分之间用点分隔。最左边的是主机名,其余部分是该主机所属的 DNS 域。因此一个 DNS 名称应该表示为“主机名+DNS 域”的形式。

要想成功部署 DNS 服务,运行 Windows Server 2016 的计算机中必须拥有一个静态 IP 地址,只有这样才能让 DNS 客户端定位 DNS 服务器。另外如果希望该 DNS 服务器能够解析 Internet 上的域名,还需保证该 DNS 服务器能正常连接至 Internet。

1. 配置 DNS 服务器过程

1）安装 DNS 服务器

在图 11.22 的“角色”列表中选择“DNS 服务器”，单击“下一步”按钮。在“添加角色和功能向导”中单击“添加功能”按钮，然后在向导对话框中单击“下一步”按钮，再在“确认安装”对话框中单击“安装”按钮。安装完成后 DNS 服务器模块将出现在服务器管理器仪表板中。

2）配置 DNS 服务器

打开“服务器管理器”的“工具”菜单，选择“DNS”菜单项，打开“DNS 管理器”窗口。DNS 管理器窗口如图 11.24 所示。

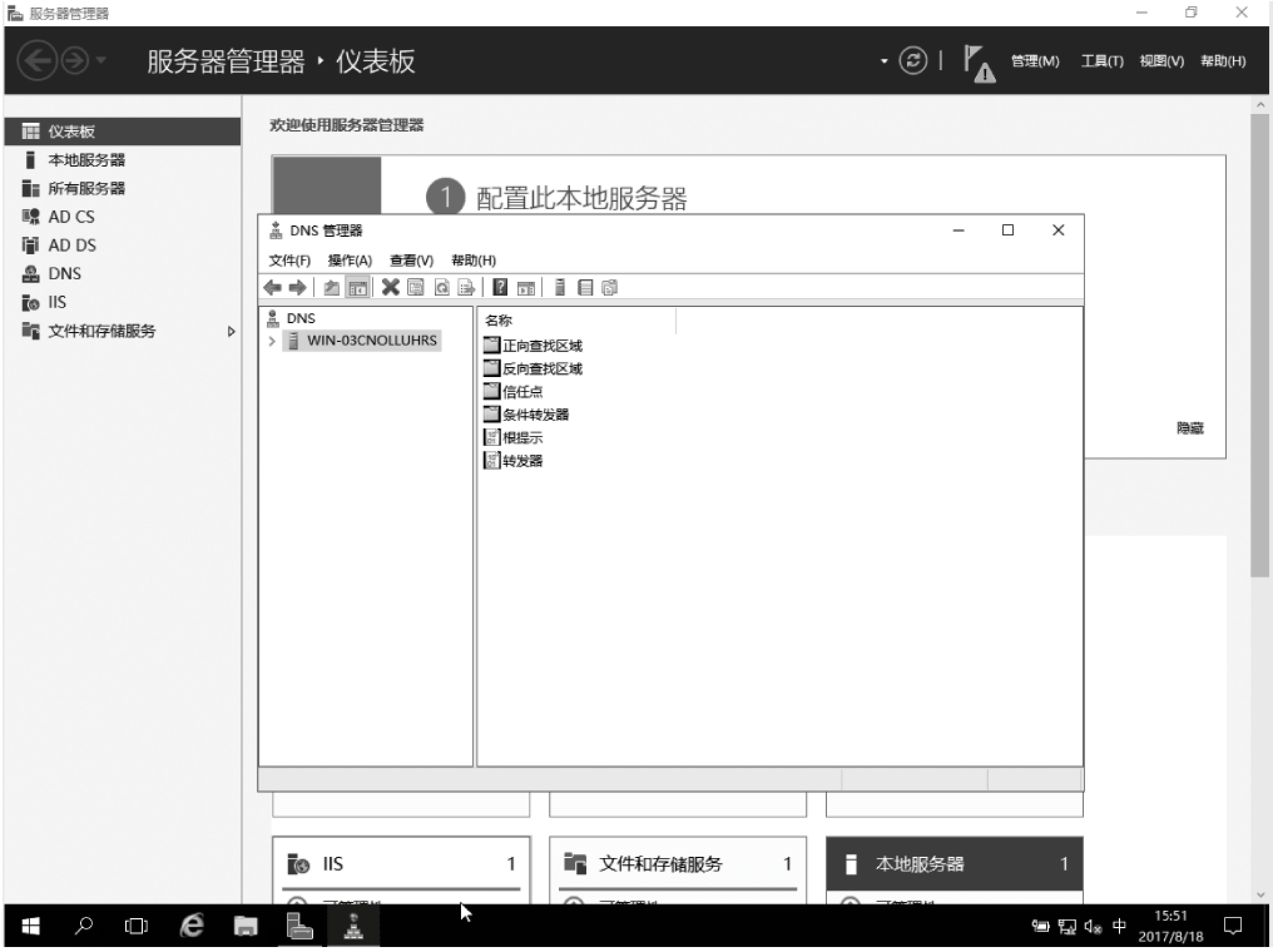


图 11.24 DNS 管理器窗口

2. 区域的类型

在 DNS 中，区域的类型有以下 3 种。

主要区域(primary zone)：用来存储此区域内所有记录的正本。当你在 DNS 服务器内建立主要区域后，可以直接在此区域内新建、修改、删除记录。区域内的记录可以存储在文件或是 Active Directory 数据库中。

辅助区域(secondary zone)：辅助区域内的每一项记录都存储在“区域文件”中，存储区域内所有记录的副本，是利用“区域复制”从其“master 服务器”复制过来的。辅助区域内的记录是只读的、不可修改的。

存根区域(stub zone)：存储着一个区域的副本信息，不过它与辅助区域不同，存根区域只包含少量记录(如 SOA、NS)，利用这些记录可以找到此区域的授权服务器。

3. 建立主要区域

DNS 客户端所提出的 DNS 查找请求，大部分是属于正向的查找(forward lookup)，也就是从主机名称来查找 IP 地址。建立步骤如下：

打开 DNS 选管理器窗口，选择 DNS 服务器，然后选择“正向查找区域”，单击“操作”菜单，在弹出的菜单中选择“新建区域”选项，在“区域类型”选择“主要区域”，在“正向或反向查找区域”中选择“正向查找区域”。输入新建区域名称，例如：zzu. edu. cn。然后在“创建新文件，文件名为”文本框中自动输入以域名为文件名的 DNS 文件。该文件的默认文件名为 zzu. edu. cn. dns(区域名＋dns)，它被保存在文件夹\windows\system32\dns 中。如果要使用区域内已有的区域文件，可先选择“使用此现存文件”一项，然后将该现存的文件复制到\windows\system32\dns 文件夹中。在动态更新选择时，选择“允许非安全和安全动态更新”选项表示任何客户端接受资源记录的动态更新，该设置存在安全隐患。选择“不允许动态更新”选项，表示不接受资源记录的动态更新，更新记录必须手动。完成后新区域“zzu. edu. cn”添加到 DNS 管理窗口。

DNS 服务器支持相当多的不同类型的资源记录，如何将几个比较常用的资源记录新建到区域内呢？

(1) 新建一项主机记录。将主机名称与 IP 地址(也就是资源记录类型为 A 的记录)新建到 DNS 服务器内的区域后，就可以让 DNS 服务器提供这台主机的 IP 地址给客户端。

(2) 新建一项主机别名。如果想要让一台主机拥有多个主机名称时，可以为该主机设置别名，例如，一台主机 host. zzu. edu. cn 当作 Web 服务器时为 www. zzu. edu. cn，而当作 FTP 服务器时为 ftp. zzu. edu. cn，但这都是同一 IP 地址的主机。

4. 建立反向查找区域

建立反向查找区域后可以让 DNS 客户端使用 IP 地址来查询主机名称。反向区域并不是必须的，可以在需要时创建。在 Windows 2003 Server 中 DNS 分布式数据库是以名称为索引而非以 IP 地址为索引。反向区域的前半部分是网络 ID(network ID)的反向书写，而后半部分必须是. inaddr. arpa。

建立一个反向查找区域与建立正向查找区域一样，步骤如下：用鼠标右键单击“反向查找区域”选项，在弹出的菜单中选择“新建区域”选项；弹出“新建区域向导”对话框，单击“下一步”；弹出“区域类型”对话框，选择“主要区域”选项，单击“下一步”；弹出“反向查找区域名称”对话框，选择“IPv4 反向查找区域”，单击“下一步”；在“网络 ID”文本框中输入正常的地址(如 211. 81. 192. 0)，这时会自动在反向查找区域名称中显示：192. 81. 211. in-addr. arpa。“区域文件”对话框的“新文件”文本框中自动输入了以反向查找区域名为文件名的 DNS 文件，192. 81. 211. in-addr. arpa. dns。“动态更新”选项和建立正向查找区域一样选择“不允许动态更新”，完成设置。反向查找区域自动添加在 DNS 管理窗口中。

11.4.4 DHCP 服务器配置

DHCP(动态主机配置协议)指的是由服务器控制一段 IP 地址范围，客户机登录服务器时就可以自动获得服务器分配的 IP 地址和子网掩码。要安装 DHCP 服务，首先，DHCP 服务器必须是一台安装有 Windows Server 系统的计算机；其次，担任 DHCP 服务器的计算机需要安装 TCP/IP 协议，并为其设置静态 IP 地址、子网掩码、默认网关等内容。默认情况

下,DHCP 作为 Windows Server 的一个服务组件不会被系统自动安装,必须添加它。

1. 安装 DHCP 服务器

在服务器管理器中启动添加服务器角色向导。在图 11.22 所示的角色选择中,选择“DHCP 服务器”,并单击“添加功能”,选择服务器所需功能,执行安装。安装完成后,DHCP 模块出现在服务器管理器仪表板的角色和服务器组中。

2. 创建 IP 地址作用域

要想为同一子网的所有客户端计算机自动分配 IP 地址,首先要做的就是创建一个 IP 地址作用域。在服务器控制器窗口中选择“工具”→“DHCP”,弹出 DHCP 管理器窗口。选择服务器→“IPv4”→“更多操作”→“新建作用域”,打开新建作用域向导。在“作用域名”步骤的“名称”文本框中为该作用域输入一个名称,如 zdxy. com,另外可以在“描述”文本框中输入一段描述性的语言,如“郑大信院”。

在“IP 地址范围”步骤,分别在“起始 IP 地址”和“结束 IP 地址”文本框中输入事先规划的 IP 地址范围的起止 IP 地址,如 192. 168. 2. 110～192. 168. 2. 200。接着在“子网掩码”文本框中输入子网掩码,或者调整子网掩码的长度值。

在“添加排除”步骤中可以指定排除的 IP 地址或 IP 地址范围,例如已经指定给服务器的静态 IP 地址需要在此排除。在“起始地址”文本框中输入准备排除的 IP 地址并单击“添加”按钮,这样可以排除一个单独的 IP 地址(也可以排除某个范围内的 IP 地址)。

在“租约期限”步骤中,默认客户端获取的 IP 地址使用期限设置为 8 天,根据实际需要修改租约期限,如 30 天。

在“路由器(默认网关)”步骤中输入网关地址,如 192. 168. 2. 1,并单击“添加”可将该路由器添加到列表。

在“域名称和 DNS 服务器”界面中可以根据实际情况设置 DNS 服务器地址。DNS 服务器地址可以设置多个,既可以是局域网内部的 DNS 服务器地址,也可以是 Internet 上的 DNS 服务器地址。

激活 IP 地址作用域,完成 DHCP 服务器配置。

3. 设置 DHCP 客户端

要想使局域网中的计算机通过 DHCP 服务器自动获得 IP 地址,则必须对客户端计算机进行相应的设置。以运行 Windows XP 系统的客户端计算机为例。

在 Windows XP 系统的桌面上右击“网上邻居”图标,在弹出的快捷菜单中选择“属性”命令。

在弹出的“网络连接”对话框中右击“本地连接”图标并在弹出的快捷菜单中选择“属性”命令,打开“本地连接属性”对话框,然后双击“Internet 协议(TCP/IP)”选项。

在弹出的“Internet 协议(TCP/IP)属性”对话框中选中“自动获得 IP 地址”和“自动获得 DNS 服务器地址”单选按钮,并单击“确定”按钮使设置生效。

进行上述操作后,DHCP 客户端的设置就正常完成。

11.4.5 FTP 服务器配置

FTP 服务器是在 Internet 上提供文件存储和访问服务的计算机,它们依照 FTP 协议提供服务。FTP 是专门用来传输文件的协议,支持 FTP 协议的服务器就是 FTP 服务器。

FTP 也是一个客户机/服务器系统，用户通过一个支持 FTP 协议的客户机程序，连接到在远程主机上的 FTP 服务器程序。用户通过客户机程序向服务器程序发出命令，服务器程序执行用户所发出的命令，并将执行的结果返回到客户机。例如，用户发出一条命令，要求服务器向用户传送某一个文件的一份副本，服务器会响应这条命令，将指定文件送至用户的机器上。客户机程序代表用户接收到这个文件，将其存放在用户目录中。

安装配置 FTP 服务器的步骤如下：

在服务器管理器中通过“添加角色和功能”加入 FTP 服务器。在安装 Web 服务器时如果已选 FTP 服务器，则可略过此步骤。在图 11.22 所示的角色选择中，展开 Web 服务器(IIS)节点，选择“DHCP 服务器”，如图 11.25 所示，并单击“添加功能”，选择服务器所需功能，执行安装。

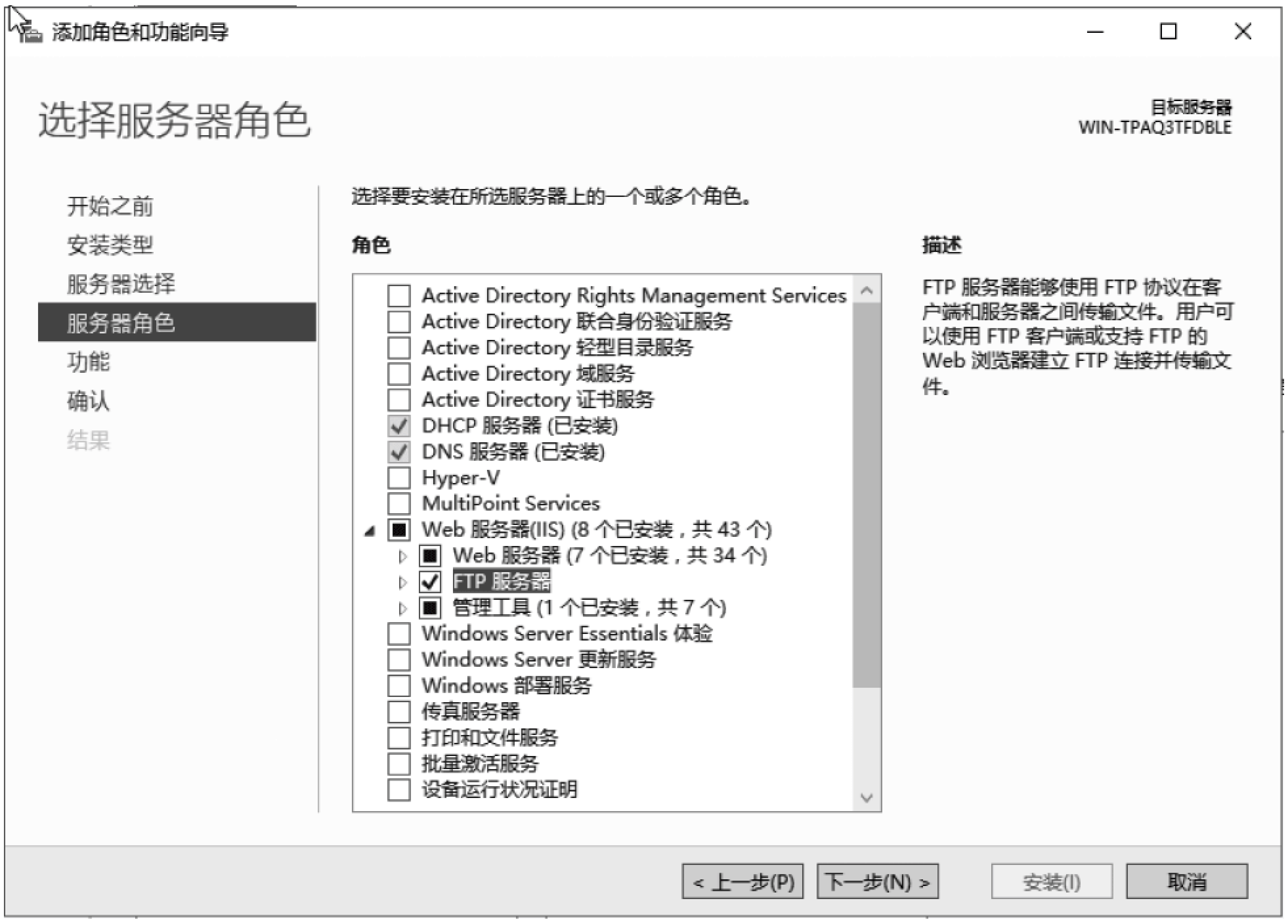


图 11.25　选择 FTP 服务器角色

习题

一、单项选择题

1. WindowsNT 是一种(　　)。

   A. 单用户多进程系统　　　B. 多用户多进程系统

   C. 单用户单进程系统　　　D. 多用户单进程系统

2. 对网络用户来说，操作系统是指(　　)。

   A. 能够运行自己应用软件的平台

   B. 提供一系列的功能、接口等工具来编写和调试程序的裸机

   C. 一个资源管理者

D. 实现数据传输和安全保证的计算机环境

3. 网络操作系统是一种(　　)。

A. 系统软件　　B. 系统硬件　　C. 应用软件　　D. 支援软件

4. 下面的操作系统中,不属于网络操作系统的是(　　)。

A. Netware　　B. UNIX　　C. Windows NT　　D. DOS

5. 下面不属于网络操作系统功能的是(　　)。

A. 支持主机与主机之间的通信

B. 各主机之间相互协作,共同完成一个任务

C. 提供多种网络服务

D. 网络资源共享

6. 用户与UNIX操作系统交换作用的界面是(　　)。

A. Windows窗口　　B. API　　C. shell　　D. GUI

7. UNIX操作系统是一种(　　)。

A. 单用户多进程系统　　B. 多用户单进程系统

C. 单用户单进程系统　　D. 多用户多进程系统

二、简答题

1. 网络操作系统(NOS)可分为哪几部分?各部门的功能是什么?
2. 简述Windows NT操作系统的特点。
3. 列举出4种Windows NT支持的网络协议,分别简要说明它们的特点。
4. UNIX操作系统通常被分为哪三个主要部分?各部分的功能是什么?
5. 什么是应用编程接口API?它是应用程序和谁的接口?
6. 试举出常用的几种系统调用的函数,说明它们的用途。

三、论述题

1. 说明网络操作系统的各项功能。
2. 网络操作系统的安全性表现在哪些方面?
3. Windows NT有哪几种域模型?它们各有什么特点?分别适用于什么环境?
4. Linux有哪些特点?

四、实验题

完成DHCP服务器的配置。

第12章 VLAN与局域网

针对局域网建设维护，如何选择合适的路由协议，实现 VLAN 间路由互连；如何依据网络的拓扑结构，选择合适设备和介质类型；如何有效利用网络资源、正确配置 IP 地址、合理划分 VLAN 等，这些都是在网络设计中需要慎重考虑的。本章将围绕有关网络问题，重点介绍 VLAN 协议、链路、划分、路由、配置以及局域网规划设计等技术。

12.1 VLAN 协议及其技术

按照一定的组网原则，将一个实际的物理网络划分成若干个小的逻辑网络，这些逻辑网络就是所谓的虚拟局域网（Virtual Local Area Network，VLAN），这种 VLAN 技术在网络建设中发挥了出强大的功能。

通过合理配置 VLAN，可以实现灵活的网络管理，同时在网络迁移时，可以轻松修改网络的设计，而不需要进行修改网络的布线等烦琐、耗时的工作。而各个 VLAN 之间是不能随意互相访问的，因为各个 VLAN 的流量实际上已经在物理上隔离开了。

12.1.1 VLAN 协议

VLAN 在以太网帧的基础上增加了 VLAN 头，用 VLAN ID 把用户划分为更小的工作组，利用 ID 限制不同工作组间的用户互访，从而控制整个网络，并简化了对网络的管理，推动了网络的发展。下面介绍的是 IEEE 802.1Q 标准，它定义了基于端口的 VLAN 模型。标准以太网帧和 802.1Q 帧格式关系如图 12.1 所示。

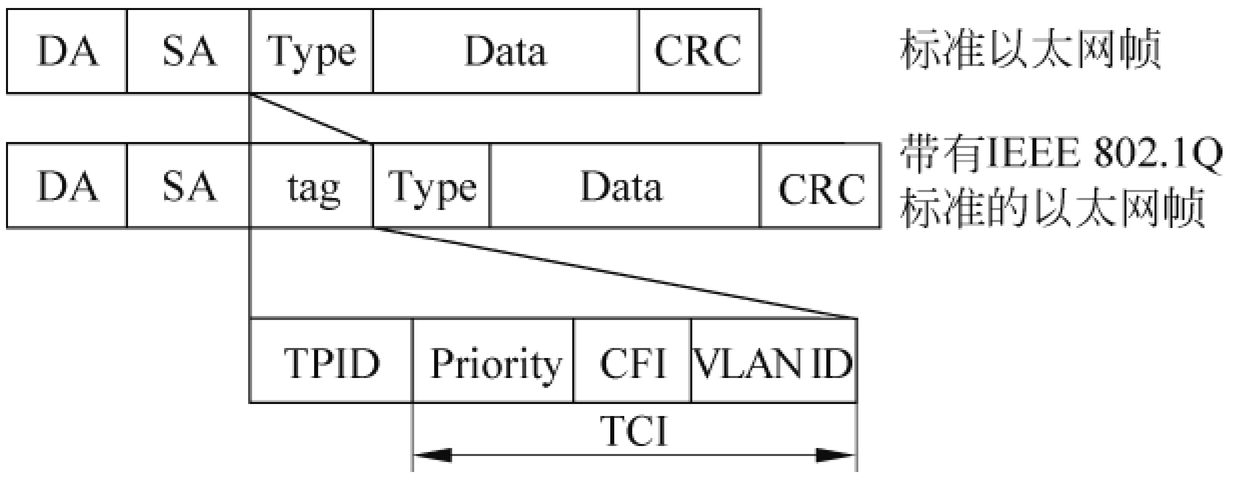


图 12.1　标准以太网帧和 802.1Q 帧格式关系

IEEE 802.1Q 定义了同一个物理链路上承载多个子网的数据流的方法和 VLAN 帧格式，为识别帧属于哪个 VLAN 提供了一个标准的方法。这个格式统一了标识 VLAN 的方法，有利于保证不同厂家设备配置的 VLAN 可以互通。IEEE 802.1Q 协议不仅规定 VLAN 中的 MAC 帧的格式，而且还制定诸如帧发送及校验，回路检测，对业务质量（QoS）参数的支持以及对网管系统的支持等方面的标准。即在标准的以太网帧中，源地址（SA）后增加一个 4 字节的 802.1Q 帧头（tag），tag 包含了 2 字节的标签协议标识（TPID）和 2 字节

的标签控制信息(TCI)。

图 12.2 给出了 IEEE 802.1Q 的标准帧格式,tag 字段可以根据其携带的 VLAN 信息,表明该数据帧属于哪个 VLAN,从而确定数据帧的属性。以下介绍 TPID 和 TCI。

目标MAC地址(DA)	源MAC地址(SA)	标记(tag)				类型(Type)	数据(data)	循环冗余校验码(CRC)
		标签协议标识(TPID)	标签控制信息(TCI)					
		0x8100	优先级(Priority)	指示符(CFI)	VLAN标识(VLAN ID)			

图 12.2　IEEE 802.1Q 标准帧格式

TPID(Tag Protocol Identifier)是 IEEE 定义的新的类型,表明这是一个加了 802.1Q 标签的帧。TPID 包含了一个固定的值 0x8100。

TCI(Tag Control Information)是帧的控制信息,它包含了下面的一些元素。

Priority:3 位,指明帧的优先级。一共有 8 种优先级,0~7。

CFI(Canonical Format Indicator):值为 0 时,说明是规范格式;1 为非规范格式。它被用在令牌环/源路由 FDDI 介质访问方法中,以指示封装帧中所带地址的位次序信息。

VLAN ID(VLAN Identified):是一个 12 位的域,指明 VLAN 的 ID,一共 4096 个,每个支持 802.1Q 协议的交换机发送出来的数据包都会包含这个域,以指明自己属于哪一个 VLAN。

以太网的帧有两种格式,有些帧是没有加上这 4 字节标志的,称为未标记的帧(untagged frame),有些帧加上了这 4 字节的标志,称为带有标记的帧(tagged frame)。

12.1.2　VLAN 链路

VLAN 接口类型在前面章节已做过介绍,这里接着介绍在 VLAN 组网中的干道链路(trunk link)和接入链路(access link),如图 12.3 所示。

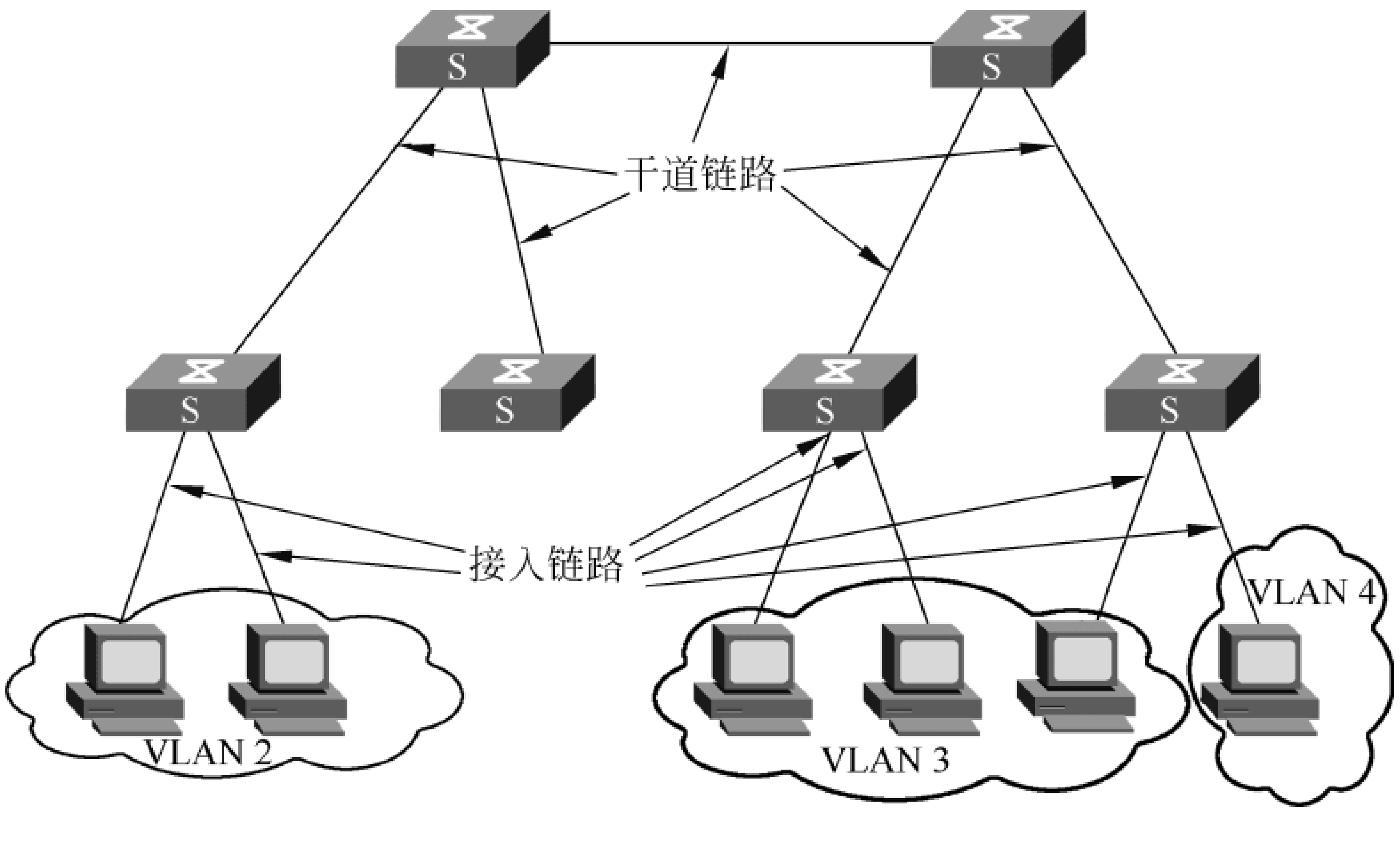


图 12.3　干道链路和接入链路

接入链路指用于连接主机和交换机的链路。通常情况下主机并不需要知道自己属于哪些 VLAN,主机的硬件也不一定支持带有 VLAN 标记的帧。主机要求发送和接收的帧都是没有打上标记的帧。接入链路属于 VLAN 某一个特定的端口,这个端口不能直接接收和发送与其他 VLAN 的通信信息。不同 VLAN 的信息必须通过三层路由处理才能转发到这个端口上。

干道链路,简称干道或干线,通常用于交换机间的互连,或者用于交换机和路由器之间的连接,它是可以承载多个不同 VLAN 数据的链路。数据帧在干道链路上传输时,交换机必须用一种方法来识别数据帧是属于哪个 VLAN 的。在 IEEE 802.1Q 的 VLAN 帧格式中可以看到,所有在干道链路上传输的帧都是打上标记的帧,交换机通过这些标记就可以确定哪些帧属于哪个 VLAN。

干道链路和接入链路不同,它不属于任何一个具体 VLAN。通过配置,干道链路可以承载所有的 VLAN 数据,也可以配置为只能传输指定的 VLAN 的数据。干道链路虽然不属于任何一个具体的 VLAN,但是可以给干道链路配置一个 pvid(port VLAN ID)。当干道链路不论因为什么原因,Trunk 链路上如出现了没有带标记的帧,交换机就给这个帧增加带有 pvid 的 VLAN 标记,然后进行处理。

对于主机来说,它不需要知道 VLAN 的存在。主机发出的报文都是 untagged 的报文;交换机接收到这样的报文之后,根据配置规则(如端口信息)判断出报文所属 VLAN 进行处理。如果报文需要通过另外一台交换机发送,则该报文必须通过干道链路传输到另外一台交换机上。当交换机最终确定报文发送端口后,将报文发送给主机之前,将 VLAN 的标记从以太网帧中删除,这样主机接收到的报文都是不带 VLAN 标记的以太网帧。

如图 12.4 所示,两台交换机通过各自的第 24 端口连接起来构成主干道,用来在两台交换机之间传送各 VLAN 的数据。交换机 A 的 VLAN1 上的主机 A 和交换机 B 的 VLAN1 上的主机 B 将可以互相通信。同样,交换机 A 的 VLAN2 上的主机 C 和交换机 B 的 VLAN2 上的主机 D 也将可以互相通信。不属于同一 VLAN 的主机 B 和主机 D,虽然同时接入交换机 B,但是无法互相通信。

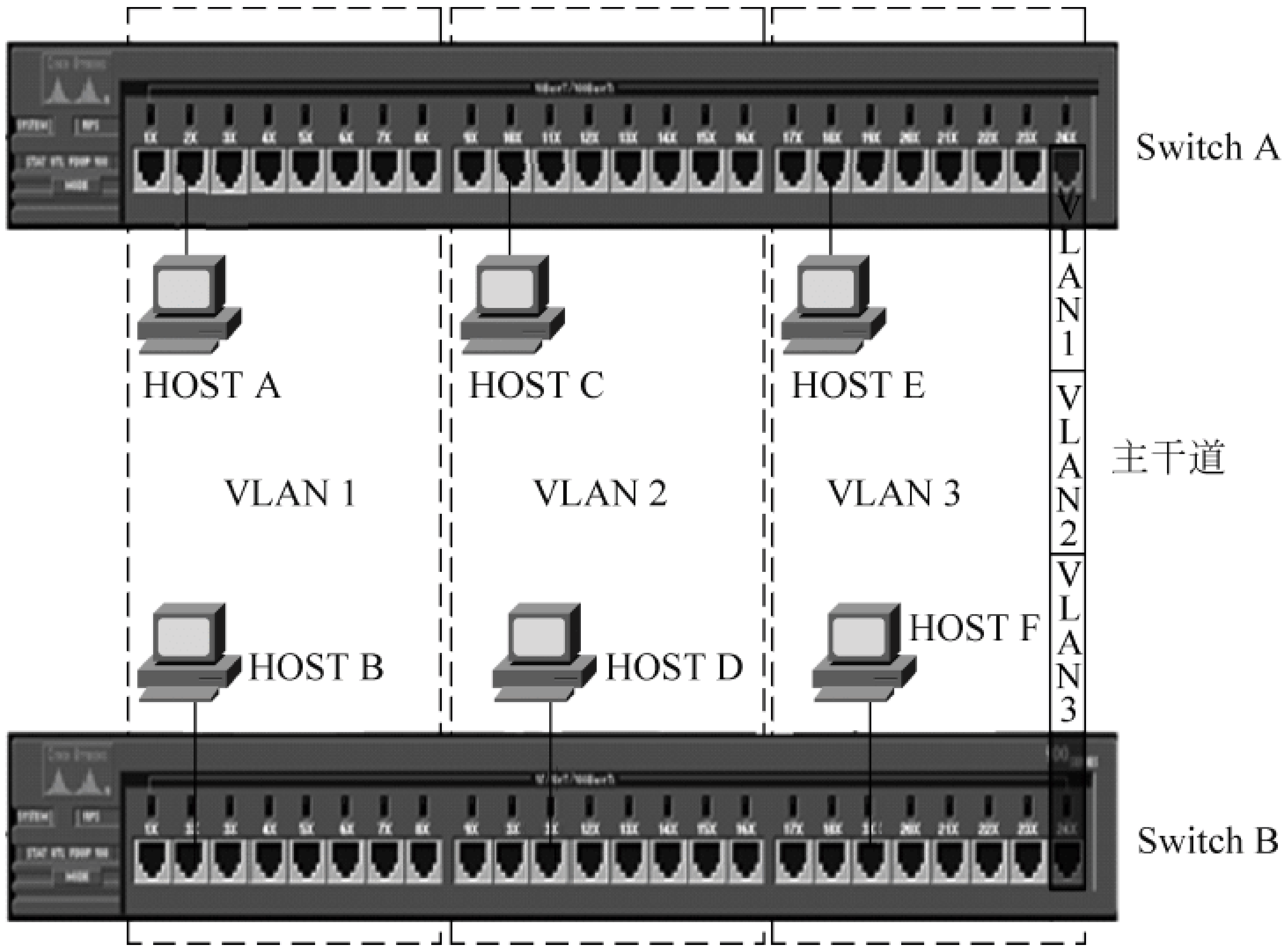


图 12.4 VLAN 干道

有两种 VLAN 协议可以选择：交换机间链路(Inter-Switch Link，ISL)和 IEEE 802.1Q。当处于动态协商模式时，交换机将视本端口对端设备类型及模式，自动协商本端口的工作模式。对端可能是普通模式，也可能是干道模式，还可能是动态模式。如果是动态模式，又可以是自动(auto)，还可以是期望(desirable)。因此，两台交换机的两个端口之间是否能够建立干道连接取决于这两个端口模式的组合，见表 12.1。

表 12.1　主干道模式组合

本方端口模式＼对方端口模式	普通	干道	自动	期望
普通	无干道	无干道	无干道	无干道
干道	无干道	干道	干道	干道
自动	无干道	干道	无干道	干道
期望	无干道	干道	干道	干道

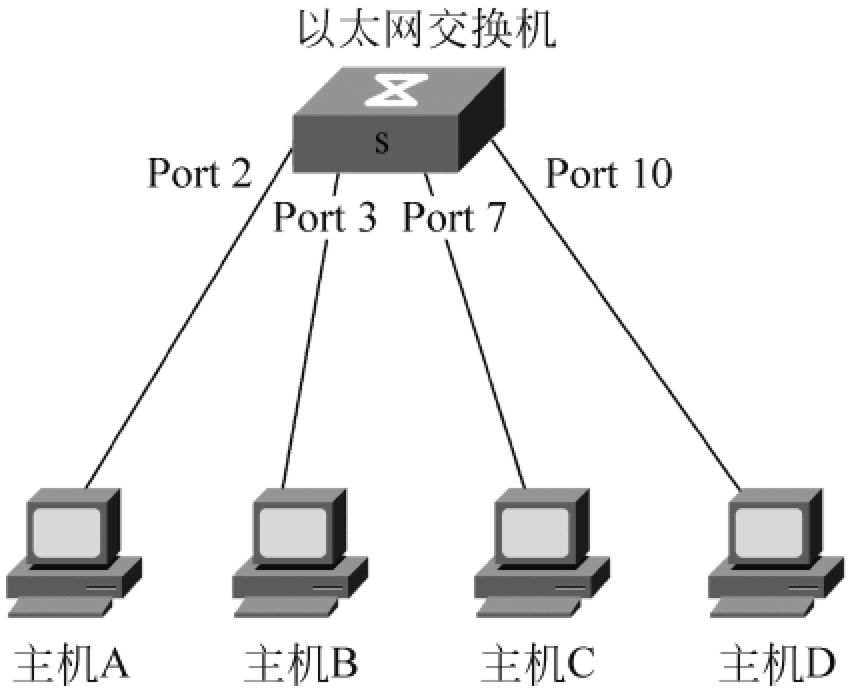
一般情况下，干道链路上传送的都是 tagged frame，接入链路上传送的都是 untagged frame。这样做的最终结果是：网络中配置的 VLAN 可以被所有的交换机正确处理，而主机不需要了解 VLAN 信息。

无论一个网络由多少个交换机构成，或 VLAN 跨越了多少个交换机，按照 VLAN 的定义，一个 VLAN 就确定了一个广播域。广播报文能够被在一个广播域中的所有主机接收到，也就是说，广播报文必须被发送到一个 VLAN 中的所有端口。

12.1.3　VLAN 划分

1. 基于端口划分 VLAN

基于端口划分的 VLAN 如图 12.5 所示，属于同一 VLAN 的端口可以连续，也可以不连续。图中端口 2 和端口 7 被指定属于 VLAN5，端口 3 和端口 11 被指定属于 VLAN10。主机 A 和主机 C 连接在端口 2、7 上，因此它们就属于 VLAN5。



VLAN表

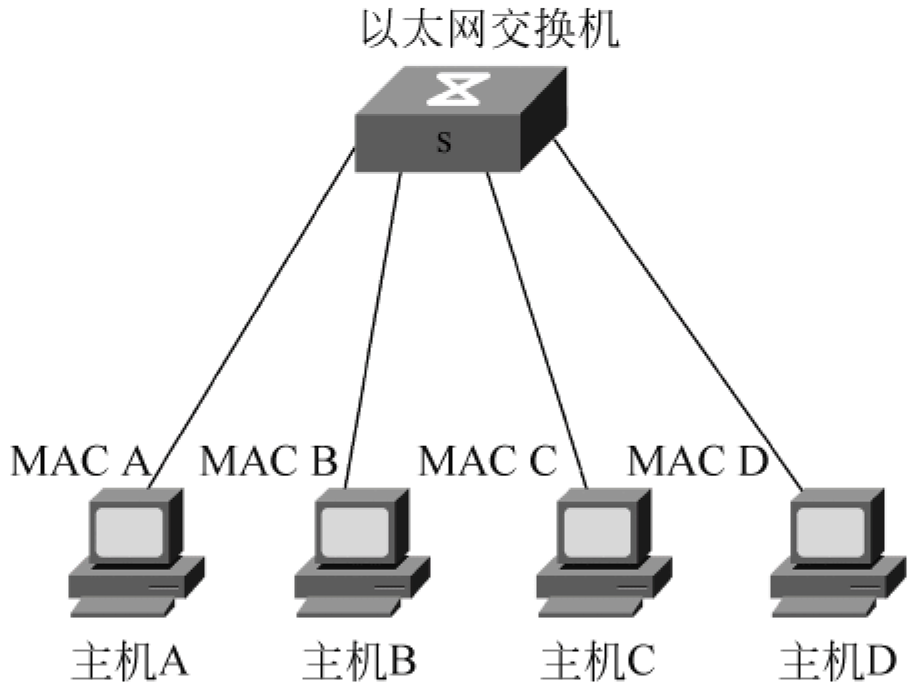
端口	所属VLAN
Port 2	VLAN5
Port 3	VLAN10
…	…
Port 7	VLAN5
…	…
Port 11	VLAN10

图 12.5　基于端口划分 VLAN

数个以太网交换机的端口也可以指定为同一 VLAN，如指定交换机 A 的 1～8 端口和交换机 B 的 1～4 端口为同一 VLAN。根据端口划分是目前定义 VLAN 的最常用的方法，其优点是定义 VLAN 成员时非常简单。

2. 基于 MAC 地址划分 VLAN

基于 MAC 地址的 VLAN 如图 12.6 所示，这种划分 VLAN 的方法是通过交换机维护一张 VLAN 映射表，这张 VLAN 表记录 MAC 地址和 VLAN 的对应关系，如 MAC A、MAC C 对应于 VLAN5。



VLAN表

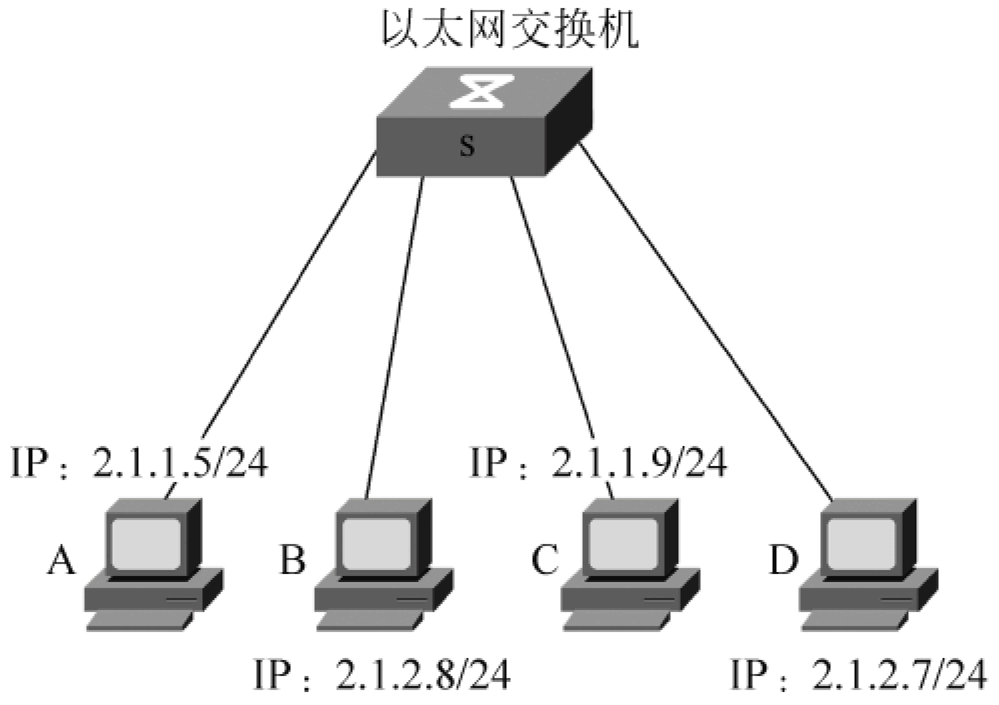
MAC地址	所属VLAN
MAC A	VLAN5
MAC B	VLAN15
MAC C	VLAN5
MAC D	VLAN15

图 12.6　基于 MAC 地址划分 VLAN

这种方法的缺点是一开始对所有的用户都必须进行配置，工作量较大。此外这种划分的方法也导致了交换机执行效率的降低，因为在每一个交换机的端口都可能存在很多个 VLAN 组的成员，这样就无法限制广播包。

3. 基于子网划分 VLAN

基于 IP 子网划分的 VLAN 如图 12.7 所示，它是根据报文中的 IP 地址决定报文属于哪个 VLAN，同一个 IP 子网的所有报文属于同一个 VLAN。这样，可以将同一个 IP 子网中的用户划分在一个 VLAN 内。图中，主机 A、主机 C 都属于 IP 子网 2.1.1.xxx，根据 VLAN 表的定义，它们因此属于 VLAN5，如果主机 C 修改自己的 IP 地址，变成 2.1.1.9，那么主机 C 就不再属于 VLAN10，而是属于 VLAN5 了。



VLAN表

IP网络	所属VLAN
IP：2.1.1.5/24	VLAN5
IP：2.1.2.8/24	VLAN10
IP：2.1.1.9/24	VLAN5
IP：2.1.2.7/24	VLAN10
…	…

图 12.7　基于子网划分 VLAN

这种方法的缺点是在检查每一个数据包的网络层地址时很费时。同时由于一个端口也可能存在多个 VLAN 的成员，对广播报文也无法有效抑制。

4. 其他 VLAN 划分方式

基于 IP 组播划分 VLAN：把一个组播划分为一个 VLAN，它不适用于局域网，通常适用于 VLAN 广域网，但是这种方法的缺点是需要有较高的处理技术和较快的处理速度。

基于规则的 VLAN：也称为基于策略的 VLAN。网络管理员可以在网络管理软件中配置相应的规则，当用户使用时，可以自动感应各种应用，并把它们划分到不同的 VLAN 中。每种网络设备可以通过不同的规则对本网中的各种应用进行划分。例如，当一个用户的数据包进入交换机中时，可以检查相应的 IP 地址，在 IP 地址-VLAN 映射表中检查相应规则，来决定属于哪个 VLAN。

按用户定义和非用户授权划分 VLAN：适用于特别的网络，在某些网络中不希望某些机器接入网络，可以在进入 VLAN 中配置相应的密码来控制接入 VLAN 的用户，这种方法的优点是安全性较高，但是实现起来仍然很麻烦。

基于协议的 VLAN：是根据二层数据帧中协议字段进行 VLAN 的划分。通过二层数据中协议字段，可以判断出上层运行的网络协议，如 IP 协议或者是 IPX 协议。如果一个物理网络中既有 IP 网络又有 IPX 等多种协议运行的时候，可以采用这种 VLAN 的划分方法。

12.1.4 VLAN 路由

局域网通常是由二层交换机构造的，每个用户都能够不受控制地直接访问网络的所有主机。在划分了 VLAN 后，通过使用路由器或三层交换机将 VLAN 互连起来的局域网，不同 VLAN 主机之间的相互通信，就需要 VLAN 间配置路由来解决。

1. 路由器提供 VLAN 间路由

通常可以认为：处于相同 VLAN 的主机叫作本地主机，本地主机之间的通信叫作本地通信；处于不同 VLAN 的主机叫作非本地主机，与非本地主机的通信叫作非本地通信。

对于本地通信，通信两端的主机同处于一个相同的广播域，两台主机之间的流量可以直接相互到达；对于非本地通信，两台主机的流量不能互相到达，主机通过 ARP 广播请求也不能请求到对方的地址，此时的通信必须借助于中间的路由器来完成，实际上是作为各个 VLAN 的网关起作用的。因此要通过路由器来互相通信的主机，必须知道该路由器的地址。在路由器配置好之后，就要在主机上配置默认网关，也就是路由器在本 VLAN 上的接口的地址。如果单独配置了路由器的地址，而没在主机上配置网关，VLAN 间的通信依然无法运行起来。

如图 12.8 所示，要执行 ping 2.2.2.20，就是主机 1.1.1.10 要同主机 2.2.2.20 通信。

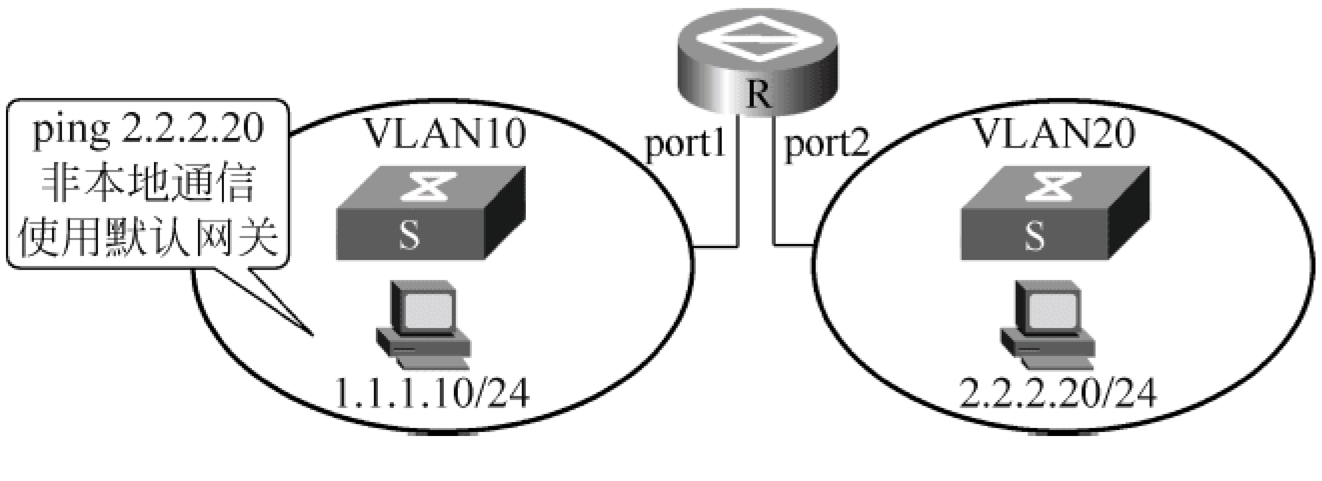


图 12.8　VLAN 间通信的路由选择

首先，主机 1.1.1.10 根据本地的子网掩码比较，发现目的主机不是本地主机，不能够直接访问目的主机；然后主机 1.1.1.10 要查找本机的路由表寻找相应的网关，在实际网络中，主机通常只配置了默认网关，因此这里主机 1.1.1.10 找到了默认网关；主机 1.1.1.10 在本机的 ARP Cache 中查找默认网关的 MAC 地址，如果没有则启动一个 ARP 请求的过

程去发现,得到默认网关的 MAC 地址后,主机将帧转发给默认网关,由路由器转发;路由器通过查找路由表,找到目的主机的 MAC 地址,将报文转发到相应的接口,发送给目的主机。

目的主机收到报文后,回应的报文经历类似的过程又转发回主机 1.1.1.10。

在这样的配置下,路由器上的路由接口和物理接口是一对一的对应关系,路由器在进行 VLAN 间路由的时候就要把报文从一个路由接口上转发到另一个路由接口上,同时也是从一个物理接口上转发到其他的物理接口上。每一个 VLAN 都要独占一个交换机端口和一个路由器的端口,这就是在传统的建网中路由器做 VLAN 间路由的局限性。

2. 使用 VLAN Trunking 技术的 VLAN 间路由

在做 VLAN 间互通的时候,对于网络中多个 VLAN,只需要共享一条物理链路。因此就出现了 VLAN Trunking(链路聚集)技术,使用 VLAN Trunking 连接如图 12.9(a)所示。

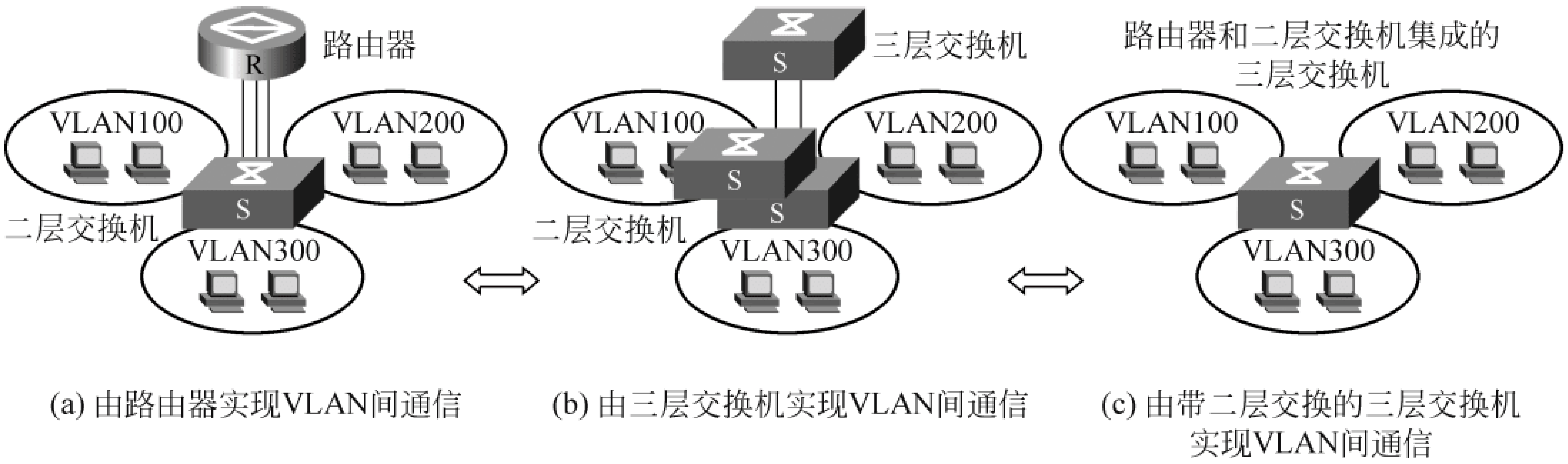


图 12.9 通过集成方式构成的三层交换机

使用 VLAN Trunking 技术网络连接,它可以使多个 VLAN 的业务流量共享相同的物理连接,是通过在物理连接上传递打标记的帧来区分各个 VLAN 流量的。

在交换机上配置连接到路由器的端口为 VLAN Trunking,在路由器上也做相同的配置。在这样的配置下,路由器上的路由接口和物理接口是多对一的对应关系,路由器在进行 VLAN 间路由的时候,把报文从一个路由接口上转发到另一个路由接口上,但从物理接口上看是在同一个物理接口的转发,而 VLAN 标记在转发后被替换为目标网络的标记。

使用 VLAN Trunking 的配置可以提高链路的带宽利用率,节省端口资源和简化管理。如当网络需要增加或删除一个 VLAN 成员时,只要操作一下设备的配置就行了。

使用 VLAN Trunking 之后,转发完全依靠软件进行,软件又要处理报文接收、校验、查找路由、选项处理、报文分片,导致性能下降,成本也高,因此诞生了大幅减少路由功能的三层交换机,如图 12.9(b)所示,在划分子网及配置 VLAN 的过程中得到了广泛的应用。

3. 基于三层和二层集成的 VLAN 间路由

目前,VLAN 间路由功能基本上都是由三层交换机来完成的。二层交换机和路由器在功能上的集成构成了三层交换机,三层交换机在功能上实现了 VLAN 的划分、VLAN 内部的二层交换机和 VLAN 间路由的功能。通过集成方式构成的三层交换机如图 12.9(c)所示。

IP 路由中绝大多数报文是不包含 IP 选项的报文,因此处理报文 IP 选项的工作在多数情况下是多余的、不同的,网络的报文长度都是不等的。为了适应不同的网络,三层交换机

采用了和路由器的最长地址掩码匹配不同的方法，使用精确地址匹配的方式处理，有利于硬件实现快速查找。三层交换机实现了简化的 IP 转发流程，利用专用的芯片实现了硬件的转发，这样绝大多数的报文处理都在硬件中实现了，只有极少数报文才需要使用软件转发，整个系统的转发性能能够得以成百倍地增加。相同性能的设备在成本上得以大幅度下降。

12.2 VLAN 配置

在前面的章节中，已经介绍过 Cisco 交换机有关 VLAN 的配置命令，下面介绍华为交换机常用配置命令，而 VLAN 间路由互连等配置实例仍以 Cisco 交换机为主。

12.2.1 VLAN 基本配置

1. 进入 VLAN 视图及创建/删除 VLAN

命令：[undo] **vlan** vlan-id

vlan-id：要进入的或要创建并进入 VLAN 的 vlan-id，取值范围为 1～4000。

2. 给 VLAN 指定端口

命令：[undo] **port** interface-list

向 VLAN 中添加或删除一个或一组端口，interface-list 通常是指接口类型、编号等。

3. 给端口指定 VLAN

命令：[undo] **port access vlan** vlan-id

把当前端口加入到指定的 VLAN 中，vlan-id 所指的 VLAN 已经存在。

例如，[Quidway-Ethernet0/1] port access vlan 3

即将 E0/1 端口加入到 VLAN3 中，VLAN3 已经存在，且 Et0/1 端口不是 Trunk 端口。

4. 设置接口为 VLAN access | trunk | hybrid

命令：**port link-type** { access | trunk | hybrid }

access：设置端口为 untagged，即为非 Trunk 端口，设为 Trunk 口后只允许 vlan1 通过；

trunk：设置端口为 tagged，即为 Trunk 端口；

hybrid：设置端口为 hybrid 端口。

5. 设置 Trunk 端口中允许通过的 VLAN Trunk

命令：**port trunk permit vlan** { vlan-id-list | all }

vlan-id-list：表示为允许通过此 Trunk 端口的 VLAN 的范围，范围是 2～4000；

all：允许所有 VLAN 通过此 Trunk 端口。

例如，[Quidway-Ethernet0/3] port trunk permit vlan 2 5 20 to 99

将 Trunk 端口 Ethernet0/3 设置为允许 2、5、20～99 等 VLAN 通过。

6. 其他命令

(1) **description** string ！string 是为了区分各个 VLAN，如小组名称、部门名称等。

(2) **display vlan** [vlan-id] ！显示 VLAN 的相关信息。

(3) **down** ！用来关闭 VLAN 接口。

(4) **up** ！打开 VLAN 接口。

12.2.2 VLAN 路由配置

1. 路由命令

1）创建/删除 VLAN 的 Router Interface 属性

命令：**interface vlan** vlan－id

vlan-id 为 VLAN 接口的标识号，只有在相应 VLAN 创建后才能创建/进入相应的 VLAN。

例如，为 VLAN100 创建路由接口：interface vlan 100

2）给 VLAN 指定创建 IP 地址和掩码

命令：**ip address** ip－address {mask | mask－length}

在相应 VLAN 接口创建后，就为其指定相应的 IP 地址和掩码。其中，

ip-address：VLAN 接口的 IP 地址；

mask：VLAN 接口 IP 地址的掩码；

mask-length：VLAN 接口 IP 地址的掩码长度。

例如，为 VLAN 2 指定 IP 地址和掩码：ip address 172.168.4.19 255.255.255.0。

2. VLAN 配置

下面将针对图 12.10，进行相关 VLAN 配置及说明。

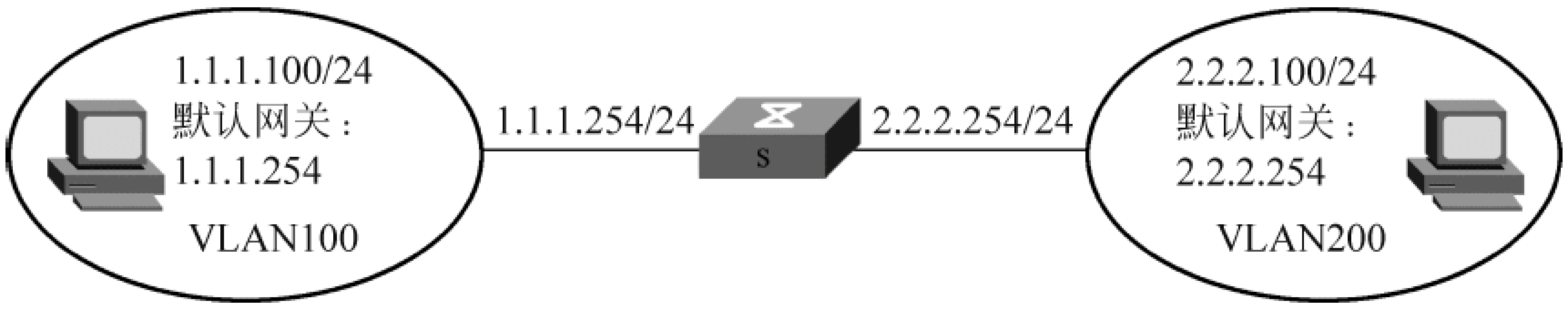


图 12.10 VLAN 配置

1）VLAN 规划端口配置

```
[Quidway]vlan 100
[Quidway－if－vlan100]port e0/1 to e0/10      !将接口 e0/1～e0/10 划归到 VLAN 100
[Quidway]vlan 200
[Quidway－if－vlan200]port e0/11 to e0/20     !将接口 e0/11～e0/20 划归到 VLAN 200
```

2）给 VLAN 配置路由接口

配置好 VLAN 后，各个 VLAN 的流量在二层上已经隔离，但仍需实现 VLAN 间的互通。在需要与其他 VLAN 通信的 VLAN 上配置三层的路由接口可以实现 VLAN 间的互通，即要配置 VLAN 接口。VLAN 接口是逻辑接口，但其意义等同于串行广域网口。

```
[Quidway]vlan 100                                 !进入 VLAN100 的配置视图
[Quidway －if－vlan100]interface vlan 100         !为 VLAN100 创建路由接口
[Quidway －if－vlan100]ip address 1.1.1.254 255.255.255.0     !给 VLAN100 接口指定 IP
[Quidway －if－vlan100]vlan 200                   !进入 VLAN200 的配置视图
[Quidway －if－vlan200]interface vlan 200         !为 VLAN200 创建路由接口
[Quidway －if－vlan200]ip address 2.2.2.254 255.255.255.0     !给 VLAN200 接口指定 IP
```

3）给主机配置默认网关

依据图 12.10，在主机上配置主机的默认网关，将默认网关指向所在 VLAN 三层交换机上的接口的地址。将 VLAN100 默认网关配置为 1.1.1.254；VLAN200 默认网关配置为 2.2.2.254。

4）使用 ping 等工具测试网络

在配置完成以上项目后，可以使用 ping 等工具测试网络是否连通。

如果能 ping 通其他 VLAN 所在接口的 IP 地址，那么说明三层路由接口已经生效，可以实现跨网段的三层互通。

5）配置路由协议

在三层交换机上可根据路由规划，配置使用路由协议以及配置路由协议参数。

配置好接口后，根据网络的路由规划，可以配置静态路由，也可以配置动态路由协议用于生成路由表。三层交换机一般只支持几个常见的路由协议。可以使用 RIP、OSPF 等动态路由协议来生成路由表，这就是在后面接着要介绍的内容。

12.3 VLAN 互连

12.3.1 RIP 实现 VLAN 互连

1. 华为交换机的互连配置

如图 12.11 所示，是由以太网交换机（华为 S3526）SwitchA、SwitchB 和 SwitchC 组成的网络，SwitchC 通过以太网接口连接到子网 202.38.165.0。SwitchA、SwitchB 的以太网接口分别连接到网络 131.109.1.0 和 129.102.0.0。以太网交换机 SwitchC 和 SwitchA、SwitchB 通过以太网络 10.24.40.0 连接到一起。由于 3 台交换机配置基本类似，下面只是列出了 SwitchA 的配置。

```
[S3526A]vlan 10                                     !创建 VLAN10
[S3526A-vlan 10]port e0/1 to e0/2                   !向 VLAN10 中添加接口 e0/1～e0/2
[S3526A]interface vlan 10                           !为 VLAN10 创建路由接口
[S3526A-if-vlan 10]ip address 131.109.1.1 255.255.255.0    !给 VLAN10 路由接口指定 IP 地址
[S3526A]vlan 20                                     !创建 VLAN20
[S3526A-vlan 20]port e0/3 to e0/4                   !向 VLAN20 中添加一组接口 e0/3～e0/4
[S3526A]interface vlan 20                           !为 VLAN20 创建路由接口
[S3526A-if-vlan 20]ip address 10.24.40.1 255.255.255.0     !给 VLAN20 接口指定 IP 地址
[S3526A]rip                                         !为 A 配置 RIP 协议
[S3526A-rip]network 10.24.40.0                      !在网段 10.24.40.0 运行 RIP
[S3526A-rip]network 131.109.1.0                     !在网段 131.109.1.0 运行 RIP
```

2. 思科交换机的互连配置

根据图 12.11，系统改用思科 S3560 交换机的网络拓扑，如图 12.12 所示，中间网段由一个普通的二层交换机 Switch0 相连，不需要做任何配置，SwitchA、SwitchB、SwitchC 配置基本类似，以下给出对 SwitchA 的配置。

```
Switch#configure terminal
Switch(config)#host SwitchA
```

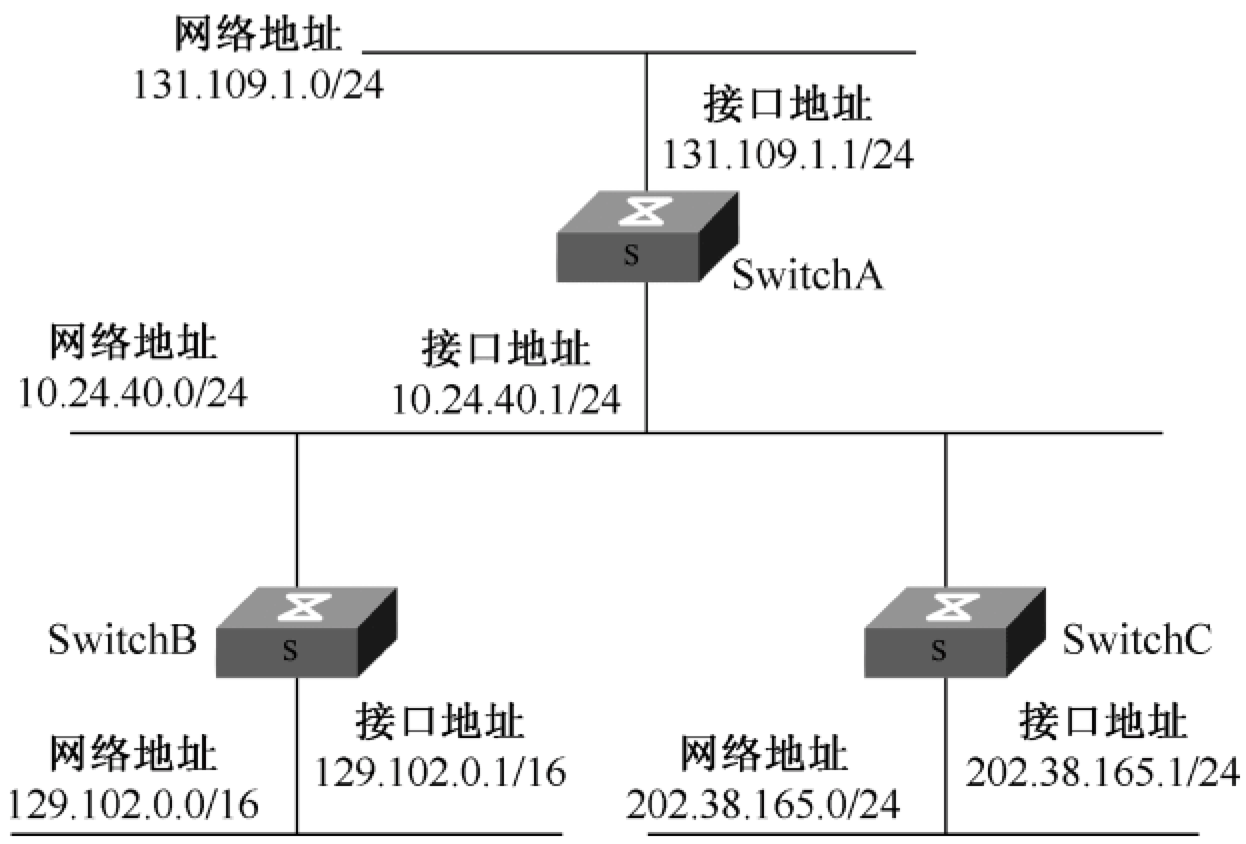


图 12.11 配置 RIP 协议

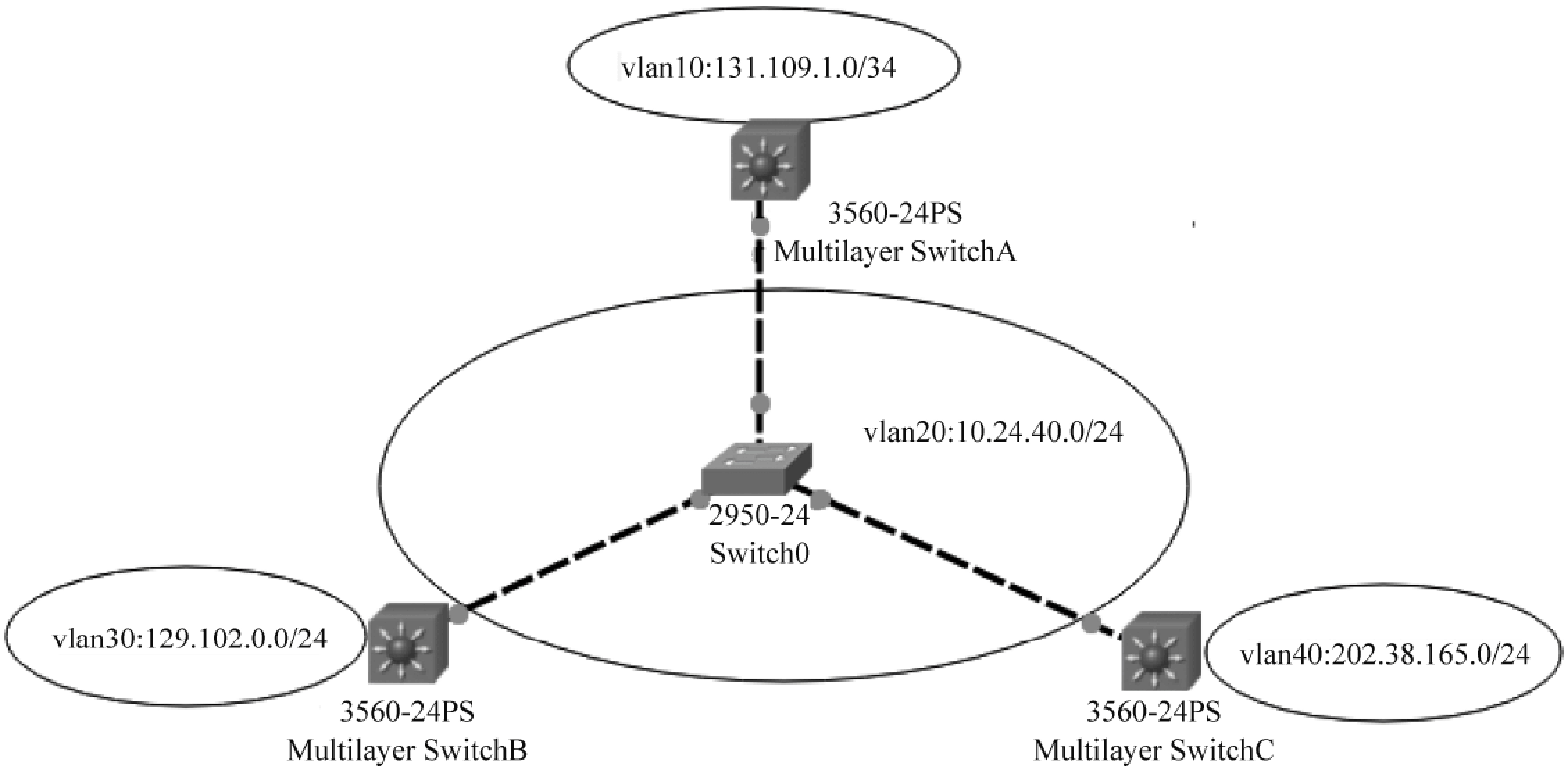


图 12.12 S3560 交换机配置 RIP 协议

```
SwitchA(config)#vlan 10                              !创建 VLAN10
SwitchA(config-vlan)#vlan 20                         !创建 VLAN20
SwitchA(config-vlan)#exit
SwitchA(config)#interface range f0/1-5               !进入一组接口 f0/1～f0/5
SwitchA(config-if-range)#switchport mode access                    !定义 f0/1～f0/5 为 access
SwitchA(config-if-range)#switchport access vlan 10
                                                     !向 VLAN10 中添加一组接口 f0/1～f0/5
SwitchA(config-if-range)#exit
SwitchA(config)#interface range f0/10-15             !向 VLAN20 中添加一组接口 f0/10～f0/15
SwitchA(config-if-range)#switchport mode access                  !定义 f0/10～f0/15 为 access
SwitchA(config-if-range)#switchport access VLAN20
SwitchA(config-if-range)#exit
SwitchA(config)#interface vlan10                     !为 VLAN10 创建路由接口
%LINK-5-CHANGED: Interface Vlan10, changed state to up          !系统报告：路由接口激活
SwitchA(config-if)#ip address 131.109.1.1 255.255.255.0         !给 VLAN10 接口指定 IP 地址
SwitchA(config-if)#interface vlan20                  !为 VLAN 20 创建路由接口
```

```
%LINK-5-CHANGED: Interface Vlan20, changed state to up          !系统报告：路由接口激活
%LINEPROTO-5-UPDOWN: Line protocol on Interface VLAN20, changed state to up
SwitchA(config-if)#ip address 10.24.40.1 255.255.255.0          !给 VLAN20 接口指定 IP 地址
SwitchA(config-if)#exit
SwitchA(config)#ip routing
SwitchA(config)#router rip                       !为 A 配置 RIP 协议
SwitchA(config-router)#version 2
SwitchA(config-router)#network 10.24.40.0        !在网段 10.24.40.0 运行 RIP
SwitchA(config-router)#network 131.109.1.0       !在网段 131.109.1.0 运行 RIP
SwitchA(config-router)#exit
SwitchA(config)#exit
SwitchA#show vlan                                !查看 VLAN 配置
SwitchA#show ip route                            !查看路由配置
SwitchA>show interface                           !查看接口配置
```

12.3.2　OSPF 实现 VLAN 互连

通过 3 个以上的交换机链状组网，划分出多个 VLAN 的组网实例，如图 12.13 所示，VLAN 间通过 OSPF 协议互连，分为 2 个区域，各交换机(SwitchA、SwitchB、SwitchC)间要设置中继路由，以下将依据拓扑图，给出网络的详细配置步骤及测试信息。

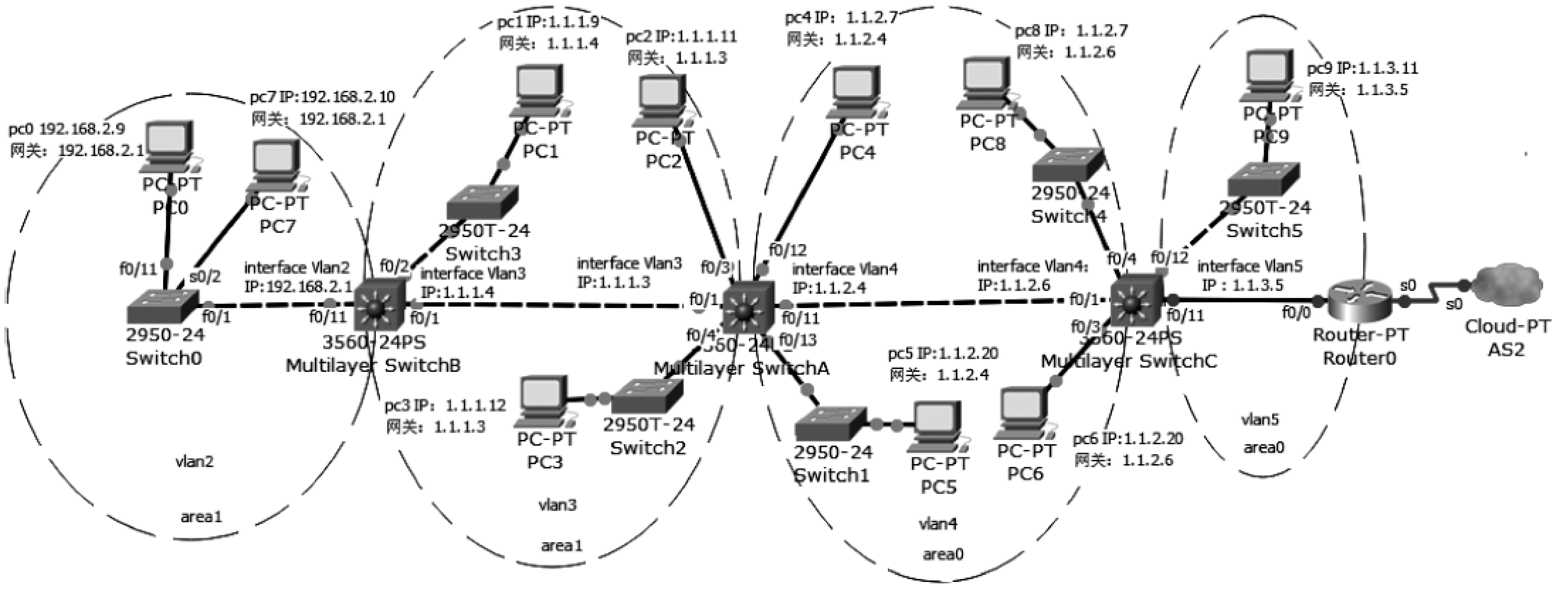


图 12.13　通过 OSPF 协议互连不同的 VLAN

1. SwitchB 配置

以太网交换机 SwitchB 接口划分为 VLAN2、VLAN3，其中继接口 f0/1 连接 SwitchA，f0/11 连接 Switch0。SwitchB 所有接口同属于一个小区，即 area1。

1）配置 SwitchB 的 VLAN2 路由接口

```
Switch>enable
Switch#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Switch(config)#vlan 2                            !开启 VLAN2
Switch(config-vlan)#exit
Switch(config)#interface range f0/11-20          !配制 VLAN2 所包含交换机的接口
Switch(config-if)#switchport mode access         !接口设置为接入模式 access
Switch(config-if)#switchport access vlan 2       !将 f0/11-20 划归至 VLAN2
```

```
Switch(config-if)#exit
Switch(config)#interface f0/11
Switch(config-if)#switchport mode trunk        !将 f0/11 由接入模式改为中继模式
Switch(config-if)#
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/11, changed state to down
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/11, changed state to up
Switch(config-if)#switchport trunk allowed vlan all        !允许所有的 VLAN 通过该中继接口
Switch(config-if)#exit
Switch(config)#interface vlan 2                !为 VLAN2 创建路由接口
Switch(config-if)#ip address 192.168.2.1 255.255.255.0        !给路由接口指定虚拟 IP 地址
Switch(config-if)#no shutdown                  !激活该路由接口
Switch(config-if)#
%LINK-5-CHANGED: Interface VLAN2, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface VLAN1, changed state to up
Switch(config-if)#exit
Switch(config)#ip routing                      !启动路由配置
Switch(config)#router ospf 1                   !启动 OSPF 协议
Switch(config-router)#network 192.168.2.0 0.0.0.255 area 1    !指定 area 1 运行 OSPF 的网段
Switch(config-router)#exit
Switch(config)#exit
```

2）配置以太网交换机 SwitchB(区域内部交换机)的 VLAN3 路由接口

```
Switch(config)#vlan 3                          !开启 VLAN3
Switch(config-vlan)#exit
Switch(config)#interface range f0/1-10         !配制 VLAN3 所包含交换机的接口
Switch(config-if)#switchport mode access       !将接口设置为接入模式
Switch(config-if)#switchport access vlan 3     !将 f0/1-10 划归至 VLAN3
Switch(config-if)#exit
Switch(config)#interface f0/1
Switch(config-if)#switchport mode trunk        !将 f0/1 由接入模式改为中继模式
Switch(config-if)#
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/1, changed state to down
%LINEPROTO-5-UPDOWN: Line protocol on Interface FastEthernet0/1, changed state to up
Switch(config-if)#switchport trunk allowed vlan all        !允许所有的 VLAN 通过该中继接口
Switch(config-if)#exit
Switch(config)#interface vlan 3                            !为 VLAN3 创建路由接口
Switch(config-if)#ip address 1.1.1.4 255.255.255.0              !给路由接口指定虚拟 IP 地址
Switch(config-if)#no shutdown
Switch(config-if)#
%LINK-5-CHANGED: Interface VLAN3, changed state to up
%LINEPROTO-5-UPDOWN: Line protocol on Interface VLAN1, changed state to up
Switch(config-if)#exit
Switch(config)#ip routing                      !启动路由配置
Switch(config)#router ospf 1                   !启动 OSPF 协议
Switch(config-router)#network network 1.1.1.0 0.0.0.255 area 1
                                               !指定 area 1 运行 OSPF 的网段
Switch(config-router)#exit
Switch(config)#exit
```

2. SwitchA 配置

以太网交换机 SwitchA 需要配置 VLAN3、VLAN4 及其路由接口，其中 f0/1 连接

SwitchB，f0/11 连接 SwitchC。SwitchA 连接 2 个不同的区域，即 area1、area0。具体配置见图 12.14。

```
Switch(config)#vlan 3
Switch(config-vlan)#exit
Switch(config)#interface range f0/1-10
Switch(config-if-range)#switchport mode access
Switch(config-if-range)#switchport access vlan3
Switch(config-if-range)#exit
Switch(config)#interface f0/1
Switch(config-if)#switchport mode trunk
Switch(config-if)#switchport trunk allowed vlan all
Switch(config-if)#exit
Switch(config)#interface vlan 3
Switch(config-if)#ip address 1.1.1.3 255.255.255.0
Switch(config-if)#no shutdow
Switch(config-if)#exit
Switch(config)#ip routing
Switch(config)#router ospf 1
Switch(config-router)#network 1.1.1.0 0.0.0.255 area 1
Switch(config-router)#exit
Switch(config)#vlan 4
Switch(config-vlan)#exit
Switch(config)#interface range f0/11-10
Switch(config-if-range)#switchport mode access
Switch(config-if-range)#switchport access vlan 4
Switch(config-if-range)#exit
Switch(config)#interface f0/11
Switch(config-if)#switchport mode trunk
Switch(config-if)#switchport trunk allowed vlan all
Switch(config-if)#exit
Switch(config)#interface vlan 4
Switch(config-if)#ip address 1.1.2.4 255.255.255.0
Switch(config-if)#exit
Switch(config)#ip routing
Switch(config)#router ospf 1
Switch(config-router)#network 1.1.2.0 0.0.0.255 area 0
Switch(config-router)#exit
Switch(config)#
```

图 12.14　以太网交换机 SwitchA 配置

3. Switch0 配置

以太网交换机 Switch0 属于二层交换机，由于和 SwitchB 相连的 f0/1 为中继模式(trunk)，这个在 SwitchB 的配置中可以发现，所以必须要进行静态配置。由于不具有路由功能，所有接口都处在 VLAN2 中，所以配置也相对简单。需要说明的是，交换机如果是以接入模式(access)与具有三层功能的以太网交换机相连时，直接连接就可以了，不必配置，除非有特殊需求，如拓扑图中的 Switch1～Switch5 就不需要再配置。以下给出的是对 Switch0 的具体配置。

```
Switch>
Switch#configure terminal
Enter configuration commands, one per line. End with CNTL/Z
Switch(config)#vlan 2
Switch(config-vlan)#exit
Switch(config)#interface f0/1                 !将与交换机连接的接口配置成中继模式
Switch(config-if)#switchport mode trunk
Switch(config-if)#interface f0/11
Switch(config-if)#switchport mode access
Switch(config-if)#switchport access vlan 2
Switch(config-if)#exit
Switch(config)#
```

4. SwitchC 运行

SwitchC 通过中继接口 f0/1 与 SwitchA 相连，通过 f0/11 与 Router0 相连，其余接口分别属于 VLAN4 和 VLAN5，都在一个 OSPF 区域(area0)。图 12.15 给出的是通过 Switch ＃show run 查看 SwitchC 的配置运行情况部分信息，具体配置与 SwitchA 相似，这里不再列出。

```
Switch#show run
!
interface FastEthernet0/1
 switchport access vlan 4
 switchport trunk encapsulation dot1q
 switchport mode trunk
!
interface FastEthernet0/11
 switchport access vlan 5
 switchport mode trunk
!
interface FastEthernet0/12
 switchport access vlan 5
 switchport mode access
!
interface Vlan1
 no ip address
 shutdown
!
interface Vlan4
 ip address 1.1.2.6 255.255.255.0
!
interface Vlan5
 ip address 1.1.3.5 255.255.255.0
!
router ospf 1
 log-adjacency-changes
 network 1.1.2.0 0.0.0.255 area 0
 network 1.1.3.0 0.0.0.255 area 0
!
```

图 12.15 查看 SwitchC 配置运行情况

5. 查看各交换机路由信息及 SwitchB 发包结果

各网络设备配置完成后，就用 show ip route 查看路由信息。图 12.16 给出的就是 SwitchA、SwitchB、SwitchC 路由表信息，如果遇到网络畅通，就可以通过路由表来查找问题。

```
Switch#show ip route
Codes: C - connected, S - static, I - IGRP, R - RIP, M - mobile, B - BGP
       D - EIGRP, EX - EIGRP external, O - OSPF, IA - OSPF inter area
       N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2
       E1 - OSPF external type 1, E2 - OSPF external type 2, E - EGP
       i - IS-IS, L1 - IS-IS level-1, L2 - IS-IS level-2, ia - IS-IS
inter area
       * - candidate default, U - per-user static route, o - ODR
       P - periodic downloaded static route

Gateway of last resort is not set

     1.0.0.0/24 is subnetted, 3 subnets
C       1.1.1.0 is directly connected, Vlan3
C       1.1.2.0 is directly connected, Vlan4
O       1.1.3.0 [110/2] via 1.1.2.6, 01:52:08, Vlan4
O    192.168.2.0/24 [110/2] via 1.1.1.4, 02:36:24, Vlan3
Switch#
```

(a) SwitchA路由表

```
Switch#show ip route
     1.0.0.0/24 is subnetted, 3 subnets
C       1.1.1.0 is directly connected, Vlan3
O IA    1.1.2.0 [110/2] via 1.1.1.3, 02:30:34, Vlan3
O IA    1.1.3.0 [110/3] via 1.1.1.3, 01:46:14, Vlan3
C    192.168.2.0/24 is directly connected, Vlan2
Switch#
```

(b) SwitchB路由表

```
Switch#show ip route
     1.0.0.0/24 is subnetted, 3 subnets
O IA    1.1.1.0 [110/2] via 1.1.2.4, 02:33:28, Vlan4
C       1.1.2.0 is directly connected, Vlan4
C       1.1.3.0 is directly connected, Vlan5
O IA 192.168.2.0/24 [110/3] via 1.1.2.4, 02:33:28, Vlan4
```

(c) SwitchC路由表

图 12.16 Switch 路由表

接下来的工作就是对 PC0～PC9 配置 IP 地址、掩码和网关地址，这里说明的是，各 PC 的 IP 地址及掩码就是 VLAN 所在子网的 IP 地址及掩码，而网关地址就是本交换机针对该 VLAN 指定的虚拟 IP 地址，也称路由接口地址。然后用 ping 命令或其他方法查看网络是否畅通无阻。图 12.17 是 SwitchB 在模拟器通过发送 PDU 的测试结果表，从中可以看到

PC7 无论到本交换机 Switch0 连接的 PC0，还是到远端 VLAN 的 PC6 都是失败的，通过查找问题，发现在 Switch0 上只配置了一个到 PC0 的接口 f0/11，并设置 f0/11 为 access 模式，属于 VLAN2，如果再添加上 PC7 所连的接口 f0/2 到 VLAN2 就可以，因此下面的测试也就全部通过了。

PDU List Window

Fire	Last Status	Source	Destination	Type	Color	Time(sec)	Periodic	Num	Edit	Delete
	Successful	PC0	PC4	ICMP		0.000	N	1	(edit)	
	Successful	PC0	PC9	ICMP		0.000	N	2	(edit)	
	Successful	PC8	PC3	ICMP		0.000	N	3	(edit)	
	Successful	PC6	PC5	ICMP		0.000	N	4	(edit)	
	Successful	PC9	PC1	ICMP		0.000	N	5	(edit)	
	Successful	PC2	PC6	ICMP		0.000	N	6	(edit)	
	Failed	PC7	PC0	ICMP		0.000	N	7	(edit)	
	Failed	PC6	PC7	ICMP		0.000	N	8	(edit)	
	Successful	PC6	PC7	ICMP		0.000	N	9	(edit)	
	Successful	PC7	PC0	ICMP		0.000	N	10	(edit)	

图 12.17 SwitchB 发包信息表测试截图

12.4 局域网设计与配置

12.4.1 局域网模型

局域网(LAN)是一个覆盖地理范围相对较小的高速容错数据网络，它包工作站、个人计算机、打印机和其他设备，为计算机用户提供了资源共享的设备访问。局域网与广域网或城域网的主要区别体现在覆盖范围、网络所有权、数据速率等方面，以及引发的实现技术。

1. 模型连接与功能

图 12.18 给出了局域网层次化设计模型连接与功能，表 12.2 列出了网络模型与各以太网之间的对应关系。局域网设计要满足如下要求：所有信息点都有交换能力；支持虚拟网划分，部门之间访问受控；网络具备容错能力，并能进行良好的网络管理；易于扩充和升级。

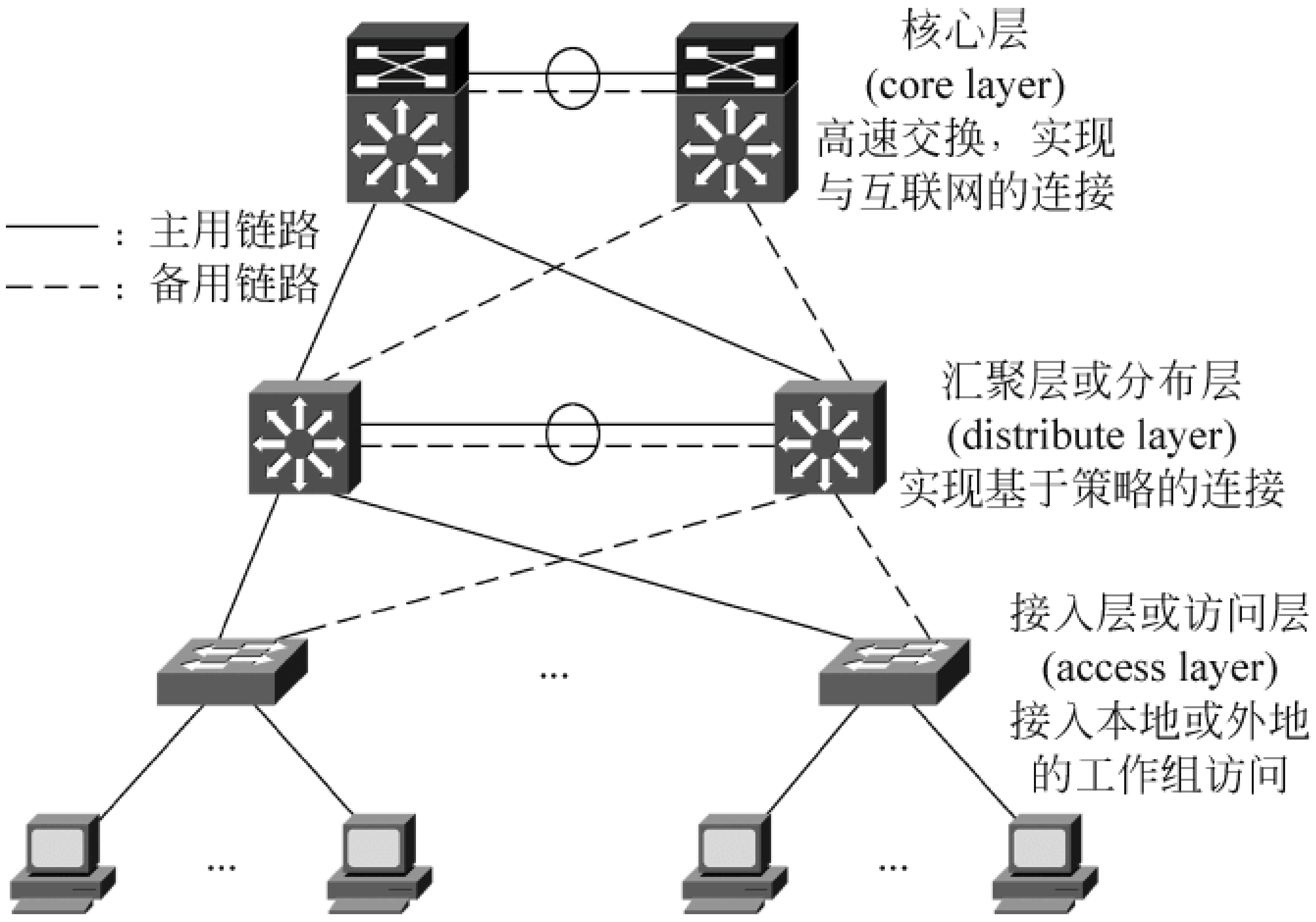


图 12.18 局域网层次化网络设计模型连接与功能

表 12.2 网络模型与各以太网之间的对应关系

模型	标准以太网(10Mbps)	快速以太网(100Mbps)	吉位以太网(1000Mbps)
接入层	终端用户与接入层交换机之间的连接	提供为高性能 PC 或工作站快速接入	通常不用
汇聚层	通常不用	提供到高速服务器,以及到其他层的快速连接	提供到高速服务器,以及到其他层的高速连接
核心层	通常不用	提供交换设备的连接	提供到汇聚层或高速服务器的高速连接,以及核心设备之间的高速互连

局域网的覆盖范围通常是一栋楼或一个园区,它的所有权和控制权也归本部门,属于内部网络,内部数据传输速率通常要高于城域网或广域网。

接入层又被称为访问层,接入层的主要功能是为网络用户提供网络接入,也可以通过访问列表或过滤进一步优化特定用户组。接入层的主要任务就是提供一个终端用户的接入点,终端用户主机通过双绞线接入该层的交换机,该层的冗余特性是通过通往其上层交换机的多条冗余链路来实现的,其安全特性可以通过基于 VLAN 成员划分来实现。

汇聚层又被称为分布层,因为网络中的所有接入层交换机在该层汇聚在一起。汇聚层是基于策略进行连接的层次,主要功能是完成网络边界的定义,其主要功能包括 VLAN 聚合、VLAN 间路由、部门或工作组级访问、广播域或多播域定义、不同类型网络介质转换和报文格式转换、实现较高的网络吞吐量、安全控制等。该层的安全控制特性可以通过制定访问控制列表的方式实现。

核心层又称为主干层,是局域网的数据交换中心,也称为局域网的主干。核心层的主要任务是提供一个高速数据包转发通道,尽可能快地完成数据的交换,也就是说各汇聚层交换机可通过核心层交换机构成的主干道进行高速的数据交换。

2. 局域网划分

(1) 根据组织的行政构成划分:组织内的行政单位是根据其功能划分的,单位内部、单位之间的数据流量和流向不同。

(2) 根据主机类型划分:局域网中的主机通常包括 PC/工作站、打印机、服务器等,其功能的差别决定了数据流量的不同。

(3) 根据主机物理分布划分:网络中主机物理位置上的分布并非总是理想和均匀的,因此网络各段的流量也不是均匀的。

(4) 根据应用类型划分:文件传输、邮件服务、公共数据库访问、共享打印等这些应用,数据流量模式不尽相同。

(5) 根据流量规则划分:80/20 规则和 20/80 规则。其中:①80/20 规则:指用户数据流量的 80%在本地网段,只有 20%的数据流量通过路由器或三层交换机进入其他网段。如果超过 20%的流量跨越网段,则会引起性能问题;②20/80 规则:由于 Internet/Intranet 应用的兴起和服务器集群的出现,使得传统的 80/20 流量模式发生了转变。网络中大部分数据流经主干,逻辑子网内部数据流量很小,用户经常需要访问本子网以外的资源,采用 20/80 规则以适应新的流量模式。

12.4.2 局域网设计

以下将以校园网为例，根据局域网设计模型来完成接入层、汇聚层和核心层的设计。

1. 校园网结构

图 12.19 给出的是校园网结构拓扑图，网络分为核心层、汇聚层和接入层。出入口路由器、防火墙等设备放置在核心层。在汇聚层将不同部门划分为不同的子网，根据不同的子网或其他要求设置 7 个虚拟局域网，分别为 VLAN10～VLAN70。VLAN 有些可以在汇聚层基于 IP 地址划分和配置，有些可以在接入层基于接口进行划分和配置。结合校园网实际规划，表 12.3 给出了局域网设计要求。

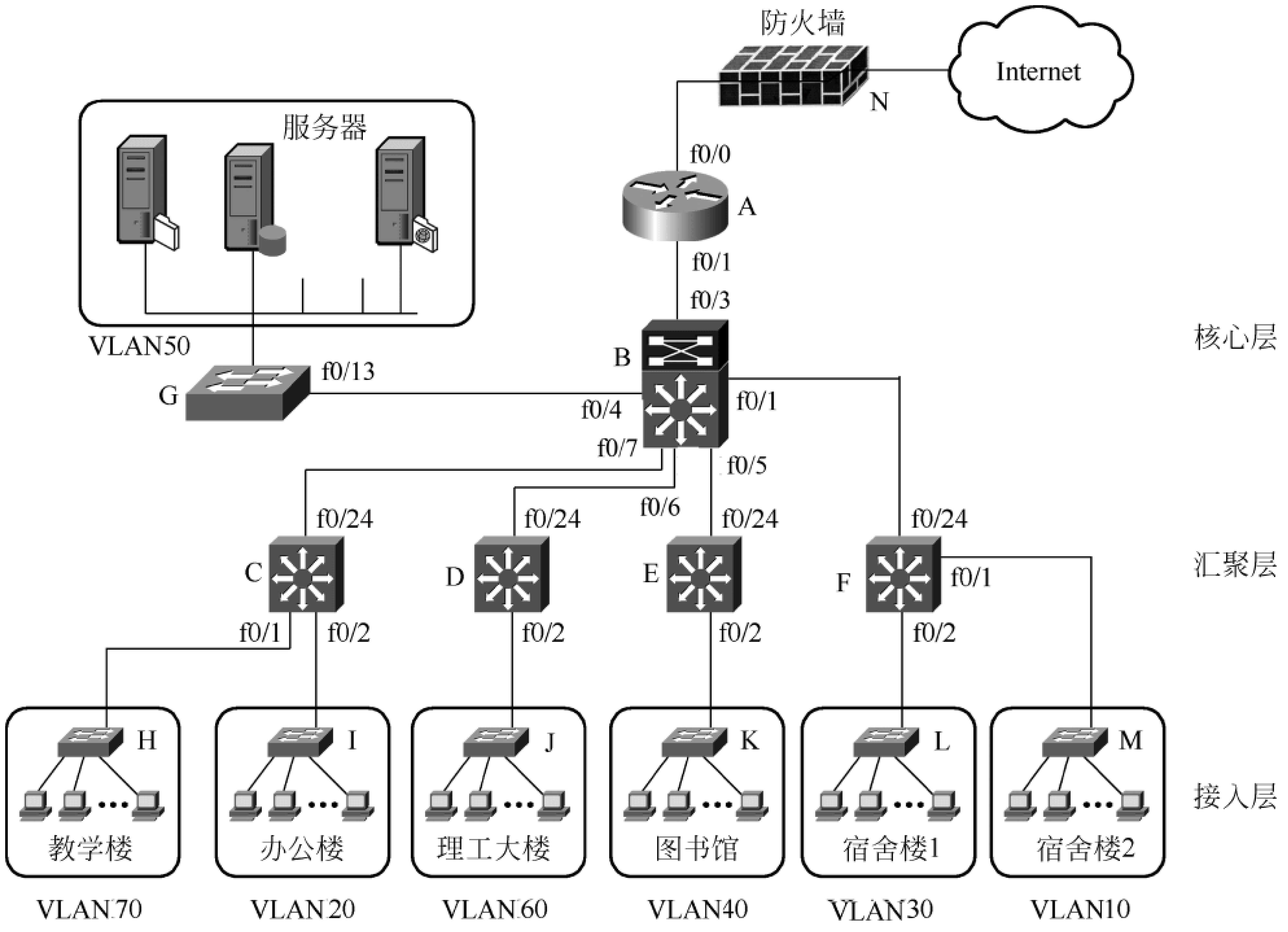


图 12.19 校园网拓扑示意

表 12.3 校园网设计要求

单位名称	规划信息点	划分 VLAN	网络设计要求
宿舍楼 1	240	VLAN10	① 信息点具有交换能力 ② 支持划分 VLAN ③ 各部门之间访问受控 ④ 能进行有效的网络管理 ⑤ 便于以后扩容或升级
宿舍楼 2	200	VLAN30	
办公楼	300	VLAN20	
图书馆	100	VLAN40	
服务器	10	VLAN50	
理工大楼	150	VLAN60	
教学楼	50	VLAN70	

2. 接入层网络

1）接入层设计要求

网络设计首先需要考虑物理链路和业务流量模型。接入层是为网络用户提供接入的层次，提供带宽，限制本地流量。要考虑降低成本，尽量提高用户带宽，管理简单方便等因素。

（1）接入层交换网络与接入普通主机连接要采用交换式以太网技术，为普通主机提供足够的带宽。

（2）宿舍楼、办公楼等信息点数较多，通过 VLAN 完成隔离和限制本地流量的功能，可以将各部门内部工作内容相同或相近的、并将共享数据的信息点划为一个工作组并安排在同一个 VLAN 中。如教学办、财务部等信息点数量少，可以不再细分 VLAN，将每个部门作为一个工作组，分配在同一个 VLAN，这些单位都可以划归于办公楼所在的 VLAN20。

（3）总校服务器经常被所有普通主机访问，要配置足够的内部链路，部门服务器相对要少一些。

（4）信息点设计规划，普通主机（PC、工作站等）的流量遵循 80/20 规则，或是 20/80 规则，或者任意。信息点中的绝大多数是连接普通主机的，要按局域网设计要求进行规划。

2）接入层设备选型

接入网络设备要选用性价比优、高接口密度的以太网交换机，并应支持网管功能。主要用的设备就是二层交换机，它的要求是低成本，有足够的接口，能提供级联接口以及高速链路接口。可以采用 Cisco 低端的交换机产品，如 Catalyst 2900、2950、3560 等。在中、大型园区网中，也可以采用高接口密度的交换机，如 Catalyst 4000 系列交换机等。

接入层选择设备为 Catalyst 2960 和 Catalyst 3560 系列交换机。

2960G-24TC-L 是一款 24 接口吉位以太网交换机，可加速网络中的桌面千兆位（GTTD）传输。主要特性有：20 个以太网 10/100/1000 接口和 4 个 10/100/1000BASE-T 双介质上行链路接口。双介质接口指的是支持两种介质接入的接口，如可接入铜缆（RJ-45）和光纤或同轴电缆（BNC），而普通以太网接口只支持 RJ-45 接入。通过高级 QoS、精确速率限制、ACL 和组播服务，实现了网络控制和带宽优化；通过多种验证方法、数据加密技术和基于用户、接口和 MAC 地址的网络准入控制，实现了网络安全性。

3560 系列交换机是一个固定配置、企业级、标准 PoE 交换机系列，能提供必要电源。它支持的应用有 IP 电话、无线接入、视频监视、建筑物管理系统和远程视频售货亭等，使用者可在整个网络范围中部署智能服务，如高级 QoS、限速、ACL、组播管理和高性能 IP 路由，且同时保持 LAN 交换的简洁性。

3. 汇聚层网络

1）汇聚层设计目标和要求

汇聚层的设计目标：足够的接口和带宽、三层和多层交换特性、灵活多样的业务能力、必需的冗余和负载平衡。

部署汇聚层：接入信息点少，汇聚层功能可以直接部署在工作组交换机上；接入信息点多，规模较大的网络要划分多个工作组，且有部门服务器，需设计汇聚层。

构建汇聚层网络：采用三层或多层交换技术，提供 VLAN 聚合和路由能力；为部门服务器提供带宽为百兆或千兆的接入；根据流量大小，确定上行链路的带宽；设计冗余链路连接到核心层，保证网络的可靠性。

通过划分 VLAN,将不同部门或工作组之间的用户隔离开;在部门一级的汇接交换机上,通过 VLAN 聚合和 VLAN 间路由功能,使这些用户可以有条件地互通和访问。

通过在接入和汇接交换机上部署冗余链路和 STP 协议,可以保证在某台汇接交换机出现故障或某条链路出现问题时,依然保证网络可用性。同样,为了保证部门服务器的可用性,也需要设置相应的冗余链路。

在汇聚层,连接服务器和核心层交换设备时,要根据接入用户需要的带宽仔细确定链路带宽。在必要时,需要考虑通过多条链路捆绑的技术来满足带宽的需要。

2) 汇聚层设备选型

汇聚层不但完成汇聚其他接入层交换机的任务,还需要为这些访问层交换机上的各个 VLAN 提供路由选择功能。因此,它必须是第 3 层的交换机,如 Catalyst 5000、6000、8500 等作为分布层交换机使用。

Catalyst 5000 交换机采用模块化的设计方式,可提供极其灵活的网络配置及极高接口密度的网络规模,是大型骨干网络的首选设备。

支持交换式/共享式 10Base-T、10Base-FL、100Base-TX、100Base-FX 以太网、吉位以太网、ATM 等。可提供最多 288 个 10Base-T 或 146 个 100Base-T、100Base-F 全双工、全交换接口。

支持第三层交换。提供 400Mbps 的交换速率,可维护 16 000 个 MAC 地址,支持 1024 个 VLAN。内部交换背板带宽 1.2Gbps,封包转发速度 1Mpps(packet per second)。

4. 核心层网络

1) 核心层设计目标和要求

核心层设计目标:足够高接口和带宽,尽可能强的数据交换能力,考虑备份和负载平衡。中小型网络的核心层功能可以同汇聚层功能合并在一台设备中,大型网络则需要分开。

在网络的核心层,凡是没有广域网连接需求,同时又需要路由器的地方,都可以用第三层交换机来代替,它的接口连接不同的子网或 VLAN。三层交换机可以被放置在小区的中心和多个小区的汇聚层,可用于网络的骨干交换机和服务器群交换机,可作为网络节点交换机。三层交换机也可以同其他以太网交换机配合使用,构造无缝的 10/100/1000Mbps 以太网交换系统,为整个信息系统提供统一的网络服务。

网络核心层主要考虑的是实现快速的数据交换,因此在考虑链路冗余以外,重点需要考虑的是物理链路的带宽。在可行的条件下,应当尽可能地部署吉比特以太网链路。

核心层的链路冗余是非常必需的,但是在考虑实施 STP 协议时需要慎重。STP 协议需要一定的收敛计算时间,对于链路或节点设备的故障反应时间会比较长——特别是对于核心层而言。在这里可以考虑采取三层路由备份协议来完成冗余链路的管理和切换功能。

2) 核心层设备选型

由于应用的需求,骨干交换机多为千兆位交换机,可以将若干条相同的源交换机与目的交换机的以太网连接线从逻辑上看成一条连接线。为了实现高速率通道以及链路冗余,可以采用核心交换机之间的吉比特以太网信道,甚至万兆位以太网信道实现主干道的连接。

网络核心层主要配备 Catalyst 3750-E 系列交换机,三层交换机属于企业级独立式可堆叠配线间交换机系列,支持安全融合、IP 语音、无线和视频等应用。通过将以太网 10/100/1000 接口和以太网供电(PoE)配置与万兆以太网上行链路相结合,完全能够满足校园网核心层的工作需要。主要特性有:设有转换器模块,可将上行链路从吉比特以太网移植到万

兆以太网；有 PoE 配置，可为所有 48 个接口提供 15.4W 功率；吞吐率高达 64Gbps；模块化电源，可带外部和可用备份电源；在硬件中提供 IPv6 路由、组播路由和访问控制列表(ACL)；带外以太网管理接口，以及 RS-232 控制台接口；支持 BGP 等。

在出口的路由器选用 Catalyst 2811，它是一款高端企业级路由器，拥有强悍的性能，能很好地为网络提供可靠的服务，并提供了安全、可扩展的网络连接，容纳了多种流量类型，如虚拟专用网络(Virtual Private Network，VPN)、动态多点 VPN(DMVPN)等。

5. 网络地址规划

学院共有图书馆、办公楼、理工大楼、教学楼和宿舍楼等组成部分，为了能够满足校园环境对 IP 地址的需求，假设采用 10.0.0.0～10.255.255.255 网段。各网段地址分配规划如下。

宿舍楼：10.2.0.0/16　VLAN10～VLAN70，其中：

宿舍楼 1 为：VLAN10，10.2.8.0/21，10.2.8.1～10.2.15.254，掩码为 255.255.248.0。

宿舍楼 2 为：VLAN30，10.2.24.0/21，10.2.24.1～10.2.31.254，掩码为 255.255.248.0。

图书馆：VLAN40，10.4.0.0/16。

服务器：VLAN50，10.5.0.0/16。

教学楼：VLAN70，10.7.0.0/16。

理工大楼：VLAN60。

表 12.4 显示的是基于接口的对网络的 VLAN 划分。在每个 VLAN 中，都要进行合适的规划，以适应网络中各种变化的需要。

表 12.4　VLAN 配置表

VLAN ID	网络地址	子网掩码	VLAN 接口地址
VLAN10	10.2.8.0	255.255.248.0	10.2.8.1
VLAN30	10.2.24.0	255.255.248.0	10.2.24.1
VLAN20	10.3.0.0	255.255.0.0	10.3.0.1
VLAN40	10.4.0.0	255.255.0.0	10.4.0.1
VLAN50	10.5.0.0	255.255.0.0	10.5.0.1
VLAN60	10.6.0.0	255.255.0.0	10.6.0.1
VLAN70	10.7.0.0	255.255.0.0	10.7.0.1

在这个网络设计中，大量采用二层交换设备，这样会出现广播风暴和存在安全性问题，所以必须要用基于接口或基于 MAC 地址的方式来划分 VLAN，后面将给出有关配置。

12.4.3 局域网配置

1. 通过 VLAN 划分配置局域网

图 12.20 是根据拓扑图 12.19 的仿真，以下主要对 SwitchF、SwitchB 进行配置。

1) 在汇聚层交换机 F 上配置 VLAN

(1) 在交换机上创建 VLAN。

```
Switch>enable
Switch#configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Switch(config)#hostname SwitchF                    !修改交换机名称
```

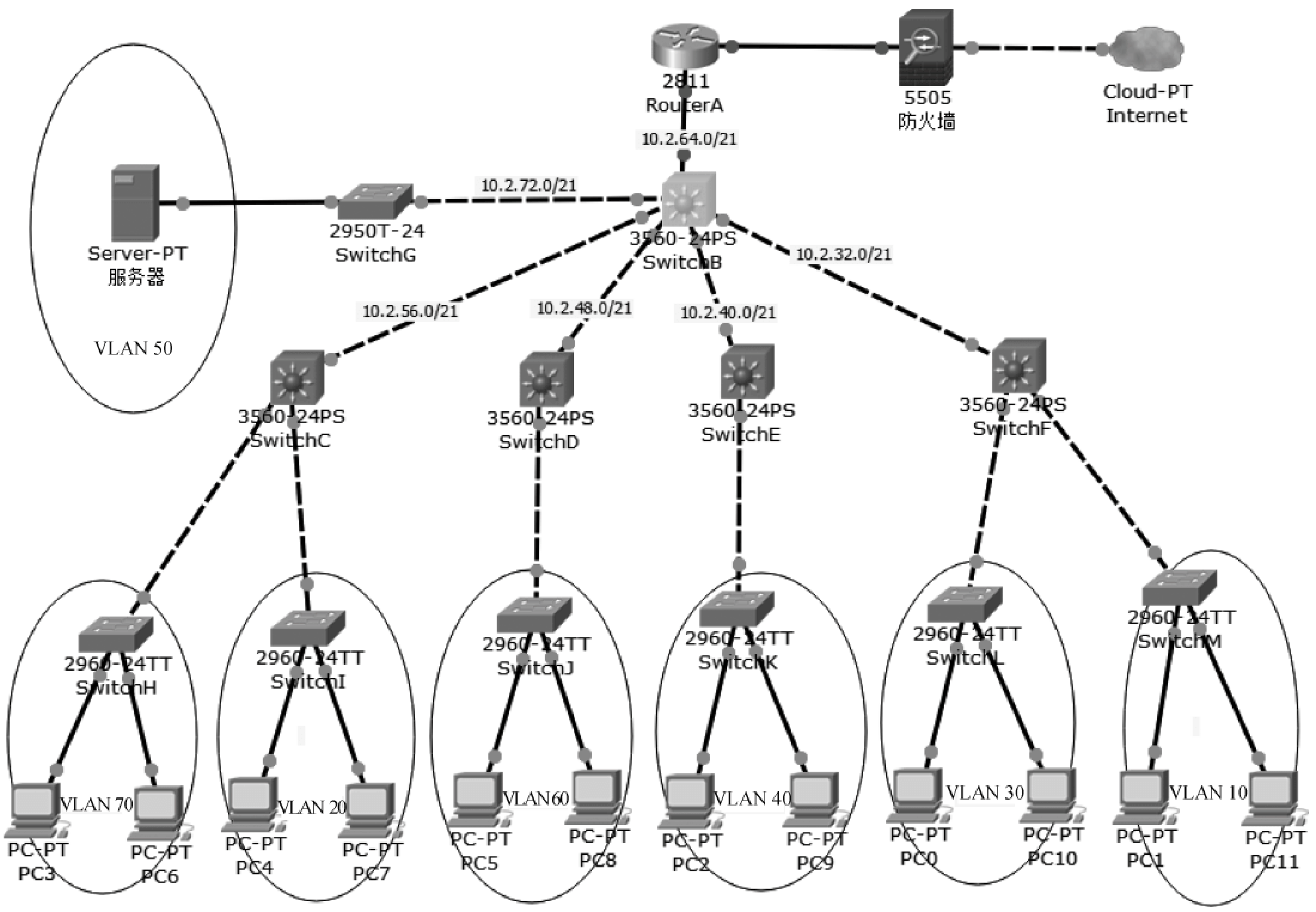


图 12.20 校园网仿真图

```
SwitchF(config)#vlan 10                        !创建 VLAN10
SwitchF(config-vlan)#vlan 30                   !创建 VLAN30
SwitchF(config-vlan)#vlan 11                   !创建 VLAN11
Switch(config-vlan)#exit
```

(2) 在交换机上将接口划分到相应的 VLAN。

```
SwitchF(config)#interface range f0/1-10        !进入 f0/1-10 接口配置
SwitchF(config-if)#switchport mode access      !接口为接入(access)模式
Switch(config-if)#switchport access vlan 10    !将当前接口加入到 VLAN10
SwitchF(config-if)#exit
SwitchF(config)#interface range f0/10-20       !进入 VLAN30 的相关接口配置
SwitchF(config-if)#switchport mode access      !接口的工作模式为 access
SwitchF(config-if)#switchport access vlan 30   !将当前接口加入到 VLAN30
Switch(config-if)#exit
```

(3) 在交换机上配置 VLAN 接口 IP 地址。

```
SwitchF(config)#interface vlan 10                        !进入 VLAN10 接口配置模式
SwitchF(config-if)#ip address 10.2.8.1 255.255.248.0     !配置 VLAN10 接口 IP 地址
SwitchF(config-if)#no shutdown                           !激活 VLAN10 接口
SwitchF(config-if)#exit
SwitchF(config)#interface vlan 30                        !进入 VLAN30 接口配置模式
SwitchF(config-if)#ip address 10.2.24.1 255.255.248.0    !配置 VLAN30 接口地址
SwitchF(config-if)#no shutdown                           !激活 VLAN30 接口
SwitchF(config-if)#exit
```

（4）指定接口成为 trunk。

```
SwitchF(config)＃interface f0/24                    !进入接口 f0/24 配置
SwitchF(config－if)＃switchport mode trunk          !f0/24 配置为 trunk 模式
Command rejected: An interface whose trunk encapsulation is "Auto" can not be configured to
"trunk" mode.                                       !命令被拒绝：此接口不能直接配置为"主干"模式
SwitchF(config－if)＃switchport mode access         !如出现以上报警时,先配置成 access 模式
SwitchF(config－if)＃switchport mode trunk          !接着再配置成 f0/24 模式就可以了
SwitchF(config－if)＃
%LINEPROTO－5－UPDOWN: Line protocol on Interface FastEthernet0/24, changed state to up
                                                    !命令报告：f0/24 改为 trunk,并被激活
SwitchF(config－if)＃switchport trunk allowed vlan all       !允许所有的 VLAN 都能通过此接口
SwitchF(config－if)＃no shutdown                 !即使出现了以上 up 状态也要配置激活,以备再启动
SwitchF(config－if)＃exit
```

（5）在交换机上中继(trunk)接口划分到相应的 VLAN。

```
SwitchF(config)＃interface vlan 11                         !进入 VLAN11 接口配置模式
SwitchF(config－if)＃ip address 10.2.32.1 255.255.248.0    !配置 VLAN11 接口 IP 地址
SwitchF(config－if)＃no shutdown                           !激活 VLAN11 接口
```

（6）配置动态路由。

```
SwitchF(config)＃ip routing                         !启动路由配置
SwitchF(config)＃router rip                         !启用 RIP 协议
SwitchF(config－router)＃version 2
SwitchF(config－router)＃network 10.2.8.0           !VLAN10 网段
SwitchF(config－router)＃network 10.2.24.0          !VLAN30 网段
SwitchF(config－router)＃network 10.2.32.0          !VLAN11 中继(主干)网段
SwitchF(config－router)＃exit
```

2）在核心交换机 B 的配置

（1）以下给出的是 SwitchB 在 SwitchF 的 trunk 连接配置，而 SwitchB 至 SwitchE、SwitchD、SwitchC、SwitchG、RouterA 的连接配置也基本类似，这里就不再赘述。

```
Switch>enable
Switch＃configure terminal
Enter configuration commands, one per line. End with CNTL/Z.
Switch(config)＃hostname SwitchB                    !修改交换机名称
SwitchB(config)＃vlan 11                            !创建 VLAN11
SwitchB(config－vlan)＃exit
SwitchB(config)＃interface f0/1
SwitchB(config)＃interface vlan 11                  !进入 VLAN11 接口配置
SwitchB(config－if)＃ip address 10.2.32.2 255.255.248.0        !配置 VLAN11 接口 IP 地址
SwitchB(config－if)＃no shutdown                    !激活 VLAN11 接口
SwitchB(config－if)＃exit
SwitchB(config)＃interface f0/1
SwitchB(config－if)＃switchport mode trunk          !将 f0/1 配置成中继 trunk 模式
SwitchB(config－if)＃
%LINEPROTO－5－UPDOWN: Line protocol on Interface FastEthernet0/1, changed state to up
```

```
SwitchB(config-if)#switchport trunk allowed vlan all
SwitchB(config-if)#exit
```

(2) 配置动态路由。

```
SwitchB(config)#ip routing
SwitchB(config)#router rip
SwitchB(config-router)#version 2
SwitchB(config-router)#network 10.2.32.0      !加入网段 SwitchB~SwitchF
SwitchB(config-router)#network 10.2.40.0      !加入网段 SwitchB~SwitchE
SwitchB(config-router)#network 10.2.48.0      !加入网段 SwitchB~SwitchD
SwitchB(config-router)#network 10.2.56.0      !加入网段 SwitchB~SwitchC
SwitchB(config-router)#network 10.2.72.0      !加入网段 SwitchB~SwitchG
SwitchB(config-router)#network 10.2.64.0      !加入网段 SwitchB~RouterA
SwitchB(config-router)#exit
SwitchB(config)#
```

(3) 在交换机 SwitchB 配置访问控制列表。

在核心交换机 SwitchB 上配置访问控制列表，可以控制各个 VLAN 之间的通信。在图 12.11 中，如 SwitchB~RouterA 的网段地址为 10.2.64.0/21；SwitchB~SwitchE 的网段地址为 10.2.40.0/21；SwitchB~SwitchF 的网段地址为 10.2.32.0/21。创建访问控制列表 9，只允许 10.2.64.0/21、10.2.40.0/21 网段与 10.2.32.0/21 之间通信，并拒绝与其他网段的数据通信。

```
SwitchB(config)#ip access-list standard 9                  !创建标准访问控制列表 9
SwitchB(config-std-nacl)#permit 10.2.64.0 0.0.7.255        !允许网段 10.2.64.0 通信
SwitchB(config-std-nacl)#permit 10.2.40.0 0.0.7.255        !允许网段 10.2.40.0 通信
SwitchB(config-std-nacl)#deny any                          !拒绝与其他网段的数据通信
SwitchB(config-std-nacl)#exit
SwitchB(config)#interface f0/1                             !进入接口 f0/1 配置模式
SwitchB(config-if)#ip access-group 9 out                   !在接口 f0/1 输出引用列表 9
```

通过以上配置，可以控制 VLAN10、VLAN30 与 VLAN40 之间通信，也可以根据实际需求对各个网段进行访问控制。

2. VTP 配置

VLAN 中继协议(VLAN Trunking Protocol，VTP)，也称为虚拟局域网干道协议，是 Cisco 的私有协议。在企业网中有比较多的交换机，配置 VLAN 工作量大，也不利于日后维护，如果每一次添加修改或删除 VLAN，都需要在所有的交换机上进行部署。因此，我们就可以使用 VTP 协议，把一台交换机配置成 VTP Server，其余交换机配置成 VTP Client，这样，有关交换机就都可以自动学习到 Server 上的 VLAN 信息。

VTP 模式有 3 种：服务器模式(Server)、客户机模式(Client)和透明模式(Transparent)。

(1) 服务器模式：提供 VTP 消息，包括 VLAN ID 和名字信息；学习相同域名的 VTP 消息；转发相同域名的 VTP 消息，以及添加、删除和更改 VLAN；将 VLAN 信息写入 NVRAM。

一个配置为 VTP Server 模式的交换机，可以向邻近的交换机广播 VLAN 配置，也可以通过它的 Trunk 从邻近的交换机学习新的 VLAN 配置。在 Server 模式下可以通过 MIB、

CLI 或者控制台模式添加、删除和修改 VLAN。

（2）客户机模式：请求 VTP 消息；学习相同域名的 VTP 消息；转发相同域名的 VTP 消息；不可以添加、删除和更改 VLAN，VLAN 信息也不会写入 NVRAM。

（3）透明模式：不提供 VTP 消息；不学习 VTP 消息；转发 VTP 消息；可以添加、删除和更改 VLAN，只将本地有效 VLAN 信息写入 NVRAM。

当交换机设置为 VTP 透明模式时，能在交换机上配置 VLAN，可使用 CLI、控制台菜单以及在使用 SNMP 管理工作站的管理信息库（Management Information Base，MIB）修改 VLAN 配置。

1）在基于 IOS 的交换机上配置

（1）创建 VTP 域命令。

```
vtp domain domain-name                    !domain-name 为 VTP 管理域名
```

（2）配置交换机的 VTP 模式。

```
vtp mode server | client | transparent
```

其中，server 为服务模式，client 为客户机模式，transparent 为透明模式。

（3）配置 VTP 口令。

```
vtp password password                     !设置密码 password
```

（4）配置 VTP 修剪。

```
vtp pruning
```

（5）配置 VTP 版本。

```
vtp ver sion 2                            !设置版本为 2，默认是版本 1
```

（6）查看 VTP 配置信息。

```
show vtp status
```

（7）其他模式配置 VTP。

如在三层交换机加一块二层挡板时，命令环境可能就会改变。比如原来需要在全局配置模式下输入 vtp 命令，而此时则需要在 vlan database 模式下输入，例如：

```
switch# vlan database                     !进入 vlan database 模式
switch(vlan) # vtp domain domain-name     !配制 VTP 管理域
```

2）配置 VTP 举例

在三层交换机下连接的二层交换机上配置不同 VLAN，实现不同 VLAN 间 PC 机的互通。把三层交换机设置成 VTP Server 模式，连接它下面的二层交换机设置成 VTP Client 模式，交换机之间的连接使用 trunk 模式，并设置为同一 vtp domain，然后开启三层交换机的路由功能即可。新出厂交换机通常的默认配置是 VLAN1，VTP 模式为服务器。具体步骤如下。

（1）将三层交换机配置成 VTP 服务器。

```
Switch# config terminal
Switch(config)# vtp mode server                 !配置 VTP Server 模式
Device mode already VTP SERVER                  !报告配置成功：已经装置为 VTP 服务器模式
Switch(config)# vtp domain abc                  !配置 VTP 域名为 abc
Changing VTP domain name from NULL to abc       !配置成功：VTP 域名变为 abc
Switch(config)# vtp version 2                   !VTP 版本 2
Switch(config)# vtp password abc123             !配置 VTP 密码为 abc123
Setting device VLAN database password to abc123 !报告完成：VLAN 数据库密码设置
Switch(config)# end
```

（2）配置 VLAN 接口。

```
Switch(config)# vlan 10                                       !进入 VLAN10
Switch(config-vlan)# name VTP pc1                             !名称为 VTP pc1
Switch(config-vlan)# vlan 20                                  !进入 VLAN20
Switch(config-vlan)# name VTP pc2                             !名称为 VTP pc2
Switch(config-vlan)# exit
Switch(config)# interface vlan 10                             !进入 VLAN10 接口配置
Switch(config-if)# ip address 192.168.10.1 255.255.255.0      !配置 VLAN10 接口地址
Switch(config-If)# no shutdown                                !激活接口
Switch(config-if)# interface vlan 20                          !进入 VLAN20 接口配置
Switch(config-if)# ip address 192.168.20.1 255.255.255.0      !配置 VLAN20 接口地址
Switch(config-if)# no shutdown                                !激活接口
Switch(config-if)# interface fa0/1                            !进入接口 fa0/1 配置
Switch(config-if)# switchport mode trunk                      !将接口 fa0/1 配置成 trunk 模式
Switch(config-if)# switchport trunk encapsulation dot1q       !trunk 封装为 dot1q 协议
Switch(config-if)# exit
```

其中，dot1q 就是 IEEE 802.1Q 协议，是 VLAN 的一种封装方式。

VLAN Trunk 有 5 种模式：on、off、desirable、auto 和 nonegotiate。

VLAN Trunk 有两种封装方式：ISL 和 IEEE 802.1Q。

另外，当 Trunk 配置后，所有 VLAN 都被加入到此 trunk 中，也可以增加或删除 trunk 中的 VLAN，但是不能删除缺省 VLAN(即 VLAN1)。

（3）将二层交换机配置成 VTP 客户机。

假设三层交换机两台二层交换机，都分别配置成 VTP Client 模式，交换机之间的连接接口配置 VLAN Trunk，以实现 VLAN 之间主机的互访。二层交换机主要配置如下。

```
Switch# config terminal
Switch(config)# vtp mode client                          !配置 VTP Client 模式
Setting device to VTP CLIENT mode.                       !设置 VTP 客户模式
Switch(config)# vtp domain abc                           !配置 VTP 域，要与 VTP Server 模式同域
Domain name already set to abc.                          !系统报告配置成功
Switch(config)# vtp version 2                            !VTP 版本 2
Switch(config-if)# interface fa0/1                       !进入接口 fa0/1 配置
Switch(config-if)# switch mode trunk                     !将接口 fa0/1 配置成 trunk 模式
Switch(config-if)# switch trunk encapsulation dot1q      !trunk 封装为 dot1q 协议
Switch(config)# end
```

3. VMPS配置

基于MAC地址的VLAN划分属于动态VLAN,主要针对TFTP服务器、VMPS数据库和交换机。VLAN管理策略服务器(VLAN Management Policy Server,VMPS)是一种基于接口MAC地址动态选择VLAN的集中化管理服务器,一般需要先创建VMPS数据库,将其保存在一个TFTP服务器上,然后才能配置和使用VMPS的管理策略。

VMPS有3种模式:open(开放)模式、secure(安全)模式、multiple(多重)模式,但用户注册工具(User Registration Tool,URT)只支持open模式。

(1) open模式。

当接口未指定VLAN时,如果该接口的MAC地址与之相关联的VLAN信息被许可,VMPS将向客户返回VLAN名。否则,VMPS将向客户返回访问被拒绝(access-denied)信息。

(2) secure模式。

当接口未指定VLAN时,如果该接口的MAC地址与之相关联的VLAN信息被许可,VMPS将向客户返回VLAN名。否则,接口将被关闭。

当接口已经指定VLAN时,如果数据库里的VLAN与MAC地址相关联的信息和接口的当前VLAN关联信息不匹配,即使配置了fallback VLAN名,接口仍将被关闭。

(3) multiple模式。

当多个MAC地址(主机)处于同一VLAN的时候,多个MAC地址可以对应一个动态接口。

例如:配置初始的VMPS服务器和备份的VMPS服务器,并设置接口为动态端口。

```
Switch(config)# vmps server 10.1.0.2 primary     !指定主 VMPS 的地址为 10.1.0.2
Switch(config)# vmps server 10.1.0.3             !指定备用 VMPS 的地址,最多可以配置 4 个 VMPS
Switch(config)# interface fa0/1
Switch(config)# switchport mode access           !设置当前接口为接入模式
Switch(config)# switchport access vlan dynamic   !设置当前接口为动态接口
```

习题

1. VLAN能起到划分广播域,提供一定安全性,分摊负载,降低延迟的作用吗?
2. 如何理解干道链路和接入链路?它们有何区别?
3. 有哪些划分VLAN的主要方式?各有何特点?
4. VLAN间是如何实现通信的?为什么说用三层交换机来连接比用路由器更好?
5. 根据图12.12所示,完成对S3560交换机SwitchB、SwitchC的配置。
6. 如何用RIP互连不同的VLAN?举例说明其配置过程。
7. 如何用OSPF互连不同的VLAN?举例说明其配置过程。
8. 局域网层次化网络设计模型分为哪几层?并说明其功能。
9. 在三层交换机下连接的二层交换机上配置不同VLAN,实现不同VLAN间PC机的互通。把三层交换机设置成VTP Server模式,连接在它下面的二层交换机设置成VTP Client模式。已知条件自行定义。